Strategic Justifications

Sanket Patil*

Job Market Paper

Click here for the latest version

January 8, 2022

Abstract

A self-interested expert obtains evidence and takes actions on behalf of many clients. Afterward, the expert justifies these actions to an auditor who has limited expertise. The auditor verifies that the expert's justification is consistent with the evidence and that the actions were in the clients' best interest. We explore how this ex-post scrutiny disciplines the expert. The constraint of justifying actions to an auditor, even an auditor with little expertise, can force the expert to act in the best interest of all clients under certain conditions. When these conditions do not hold, the expert devises a justification that makes the expert's selfish actions appear client-optimal. In this justification, the expert inflates the strength of weak evidence and deflates the strength of strong evidence. Moreover, an increase in the auditor's expertise can reduce clients' aggregate payoff.

Keywords: Justifications, Asymmetric expertise, Fiduciary duty.

JEL classification: D82, D83, D91.

*Kellogg School of Management, Northwestern University. 2211 Campus Drive, Evanston, IL 60208, USA. Email: sanket.patil@kellogg.northwestern.edu. I am indebted to my advisor Yuval Salant for his constant guidance, encouragement, and thoughtful comments. I greatly benefited from the insightful feedback and timely encouragement of my committee members Niko Matouschek and Alvaro Sandroni. I am also thankful to Daniel Barron, Henrique Castro-Pires, Peter Klibanoff, Joshua Mollner, Michael Powell, Luis Rayo, Lorenzo Stanca, Ruozhou Yan, and seminar participants at Northwestern University for helpful comments and discussions. Preliminary and Incomplete. All remaining errors are my own.

1 Introduction

This paper is about self-interested experts. In day-to-day life, experts act on behalf of clients. For example, surgeons make medical decisions for patients, and financial advisors choose investments for investors. Even when all the data essential for making decisions is available to clients, the expert possesses knowledge and models that allow the expert to precisely interpret this data and make better decisions. For example, patients require a surgeon to interpret medical records and assess the need for surgery. Investors need a financial advisor to interpret financial data and choose investments. This inability of clients to interpret commonly observable data can reduce the effectiveness of outcome-dependent contractual and reputational tools used for disciplining experts.

In the absence of such disciplining tools, a self-interested expert may be tempted to further her own interest at the clients' expense. However, the expert in these settings is often legally or ethically required to act in the clients' best interests. This institutional constraint, usually called a "fiduciary duty," makes the expert answerable to auditors and regulators who safeguard clients' interests. As we discuss later in the introduction, the auditor may still lack the expertise in interpreting the data compared to the expert. However, the auditor observes the expert's behavior across a spectrum of clients and can use it to scrutinize the expert's decisions. If the auditor detects that the expert advanced her own interests instead of clients', he can use extreme measures to punish the expert. Thus, we see medical boards revoking the license of surgeons who fail to justify their medical choices and the Securities and Exchange Commission (SEC) charging financial advisors for the breach of fiduciary duty.[1] Because of such severe consequences, it becomes imperative for the expert to appear selfless to the auditor.

Motivated by these observations, this paper considers a self-interested expert who takes binary decisions for a unit mass of clients. After taking decisions for clients, the expert must justify these decisions to an auditor with limited expertise. We study if this ex-post requirement of justifying decisions to an auditor can discipline the expert to act in the clients' best interest.

The auditor in the model observes all the data that the expert uses in the decision-making. The auditor, however, has limited expertise in interpreting this commonly observable data. Such asymmetry in the expertise of the auditor and expert may exist due to various factors. For exam-

[1]An example of the cancellation of a medical license for unnecessary intervention can be found at https://www.baltimoresun.com/news/bs-xpm-2011-07-13-ms-md-midei-license-revoked-20110713-story.html. For an example of SEC charges filed for not acting in clients' best interest, see https://www.sec.gov/news/press-release/2018-137.

ple, the auditor usually operates in time and resource-constrained environments. He has to audit several experts who usually have stronger incentives and better resources to keep up with the latest technological advances than the auditor. To capture this asymmetry, we consider an auditor who does not know the distribution of data in one of the two states of the world. However, the auditor understands that this missing piece from his knowledge must satisfy a monotonicity property and must integrate across all data realizations to one. A "justification" is then modeled as this missing piece that the expert reports to the auditor to rationalize her decisions in the data. Even though the auditor cannot verify the expert's report, he can check if the offered distribution satisfies the above-mentioned consistency properties. Thus, focusing on asymmetric expertise allows us to propose a tractable framework where one player offers a "justification" for commonly observable data, and the other player evaluates its "coherence."

The interaction between players has two stages. In the first stage, the expert acts on behalf of a continuum of clients. Each client is associated with an unknown binary state of the world, i.e., a high or a low state. The expert's task is to diagnose the unknown state of each client and take a matching action. If the expert takes high action for a client in the low state, a type-I error occurs (for example, the surgeon performs an unnecessary surgery). In this case, the client gets a payoff $-c$. The unit mass of clients differ in this payoff parameter c and are distributed uniformly on $[0, \overline{c}]$. On the other hand, the expert gets a larger payoff $-c + b$, where $b > 0$ measures the conflict of interest between the client c and the expert. The interests of the expert and a client are perfectly aligned for other combinations of the client's state and the expert's action. To diagnose the unknown state associated with client c, the expert obtains evidence $x_c \in [0, 1]$. The informational content of this evidence is summarized by the likelihood ratio $L(x_c) = f(x_c|h)/f(x_c|l)$, where $f(x_c|\omega)$ is the density function in the state ω. Function L is non-decreasing, implying that a higher realization of evidence x_c is more indicative of the high state of the world. As we show later in the paper, it is without loss of generality to assume that evidence is distributed uniformly in the low state of the world. Thus, likelihood ratio is a non-decreasing function $L(x_c) = f(x_c|h)$ that integrates across all evidence realizations to one.

A *decision rule* summarizes the expert's behavior in the first stage and specifies the chance that the expert takes the high action for any given client and evidence pair. The conflict of interest between the expert and any given client implies that an unrestricted expert's threshold evidence for the high action is smaller than that of the client's.

In the second stage, the expert must justify her decisions to an auditor who observes all her decisions and all the evidence used while making them. The auditor, however, has limited expertise in interpreting this data. This is because the auditor does not know the distribution $L(x_c) = f(x_c|h)$ of the evidence in the high state of the world. However, the auditor knows that L is non-decreasing and integrates to one across all evidence realizations. The act of justifying decisions requires the expert to provide a likelihood ratio function J (*justification*) that makes her decisions in the data appear client-optimal. The justification J, however, cannot be arbitrary. It must be *feasible*, i.e., it has to be non-decreasing and must integrate to one.[2]

As long as the expert can provide a feasible justification that rationalizes her decisions in the data, she passes the auditor's scrutiny and avoids severe consequences. A *justifiable decision rule* guarantees that for any plausible realization of data, the expert can always come up with a feasible justification. We study the expert's behavior when she optimally chooses a justifiable decision rule that maximizes her payoff from interaction with clients.

The first step in solving the expert's problem is understanding the restrictions imposed by a justifiable decision rule on the expert's behavior. A technical result shows that when the expert uses a justifiable decision rule d, she can provide the same feasible justification J_d for every realization of the data and J_d completely characterizes her behavior. Consequently, for any evidence realization x, the expert takes the high action for clients with payoff parameter in $[0, \min\{J_d(x), \overline{c}\}]$. Otherwise, she takes the low action. Panel (a) in Figure 1 depicts this behavior for evidence realizations x_1 and x_2. Using this result, we transform the expert's problem of finding the optimal justifiable decision rule into a problem of identifying an optimal feasible justification J^*. We term it *strategic justification*.

To understand the trade-off faced by the expert in determining the strategic justification J^*, fix an evidence realization x. The auditor wants the expert to take high action for clients with payoff parameter in $[0, \min\{L(x), \overline{c}\}]$. The expert, on the other hand, prefers to take the high action for a larger set of clients $[0, \min\{L(x)+b, \overline{c}\}]$. If $L(x) < \overline{c}$, then conflict between the expert and auditor is given by the set of clients $(L(x), \min\{L(x)+b, \overline{c}\}]$. To further her own interest at evidence realization x, the expert wants to take high action for as many clients as possible in this interval. However, to make these selfish actions appear client-optimal, the expert must appropriately inflate $J^*(x)$ above the actual likelihood ratio $L(x)$. Consider the case where the actual likelihood ratio function is

[2] For most of the paper, the auditor does not draw inference from the choice of the expert's justification and deems a justification feasible as long as it satisfies specific consistency properties. We later show that our analysis and insights extend even when the auditor reasons strategically and uses a more stringent criterion to assess the feasibility of justifications.

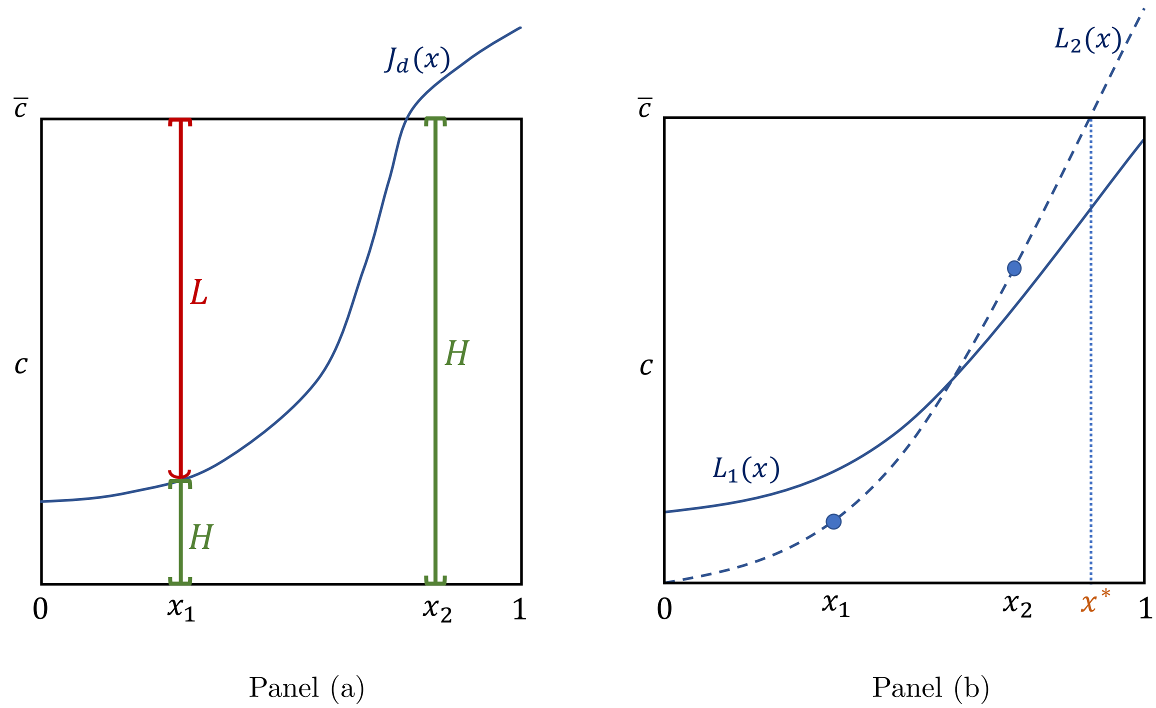


Figure 1: Graphical illustration of intuition (Evidence realization is plotted on the horizontal axis, while the vertical axis has payoff parameter c.)

such that the condition $L(x) < \bar{c}$ holds for all evidence realizations (for example, L_1 in Panel (b) of Figure 1). In this case, the expert wants to inflate J^* at all evidence realizations. However, such global inflation is not feasible because the resulting justification does not integrate to one and may lead to non-justifiable decisions. Thus, the expert must decide on which evidence realizations and how much to inflate J^*. In this case, we show that the expert's strategic justification J^* coincides with the actual likelihood ratio function L almost everywhere. The expert, therefore, acts in the best interest of all clients.

Suppose that $L(x) \geq \bar{c}$ for evidence realizations $[x^*, 1]$, where $0 < x^* < 1$ (for example, L_2 in Panel (b) of Figure 1). For these evidence realizations, there is no conflict between the expert and the auditor. This is because both of them want the high action for all clients $[0, \bar{c}]$. Thus, the conflict of interest between the expert and the auditor exists only on evidence realizations $[0, x^*)$. In this case, we establish that there exists an evidence realization $x_T \geq x^*$ such that the expert's strategic justification $J^*(x)$ is above the actual likelihood ratio $L(x)$ for evidence realizations smaller than x_T, and otherwise, the order is reversed. This strategic justification allows the expert to take selfish decisions and mask them as client-optimal.

In situations where the expert can successfully justify her selfish decisions as client-optimal, we study the role of the auditor's expertise in improving clients' payoffs. Contrary to the expectation, we show that more expertise of the auditor can reduce clients' aggregate payoff. To facilitate this comparative statics exercise, we consider an auditor with more expertise than just knowing a few consistency properties of the evidence-generating process. For example, the surgeon on the medical board may have had experience interpreting certain symptoms and test results. Such additional expertise of the auditor is modeled through a finite set of *anecdotes*. Formally, anecdotes are evidence realizations where the auditor knows the value of L precisely, and therefore, has the same expertise as the expert in interpreting these evidence realizations. However, for the remaining evidence realizations, the auditor can infer L only imperfectly. Panel (b) of Figure 1 illustrates set $\{x_1, x_2\}$ of anecdotes when the actual likelihood ratio function is L_2.

We identify conditions under which providing an auditor with more anecdotes can reduce clients' aggregate payoff. The direct effect of an additional anecdote is that it restricts the expert's ability to inflate the strategic justification J^* at evidence realizations in the vicinity of the anecdote. The indirect effect comes through a change in the expert's strategic justification at evidence realizations that are not in the vicinity of the anecdote. At these evidence realizations, J^* may move away from L, causing harm to the clients. This indirect effect dominates the direct effect and reduces the clients' aggregate payoff.

Literature review: Our paper relates to the literature on mechanism design and particularly to the literature on delegation (Holmstrom (1978), Melumad and Shibano (1991), Alonso and Matouschek (2008)). The principal in these papers decides how much discretion to give to an informed agent. Such settings assume that the principal has the same expertise but observes fewer data. In contrast, our paper focuses on settings with asymmetric expertise between the principal (client) and the agent (expert), where it is as if the expert has all the discretion. In such settings, we show that the institutional constraint of "fiduciary duty," modeled as ex-post justification, can discipline a self-interested expert.[3]

The mechanism through which the expert is disciplined in our paper is similar to the "quota" mechanisms of Jackson and Sonnenschein (2007), Frankel (2014, 2016). These mechanisms align the preferences of the principal and the agent when the agent privately observes data drawn from a commonly known distribution. Our approach identifies a similar mechanism that works in the presence of asymmetric expertise and commonly observed data. However, for this mechanism to

[3] A recent paper by Bhattacharya, Illanes and Padi (2020) empirically investigates the effectiveness of fiduciary duty in improving the quality of financial advice.

work, only specific properties of the data generating process rather than the complete details need to be common knowledge among the players.

Spiegler (2002) is an early contribution that studies the effects of ex-post justifications on strategic interactions. A critic with preferences identical to the player finds the player's strategies justifiable if the player can answer his criticism. In contrast, the auditor in our paper acts based on different preferences than the expert's and deems a justification feasible if it satisfies specific consistency properties.[4]

The literature on information transmission asks if and how the expert can influence decisions of less-informed players (Crawford and Sobel (1982), Milgrom (1981), Grossman (1981), Kamenica and Gentzkow (2011), and Rayo and Segal (2010)). In this large literature, our paper relates to papers on boundedly rational persuasion. Eliaz, Spiegler and Thysen (2021) consider a sender who persuades a receiver by providing messages as well as their interpretations. In contrast, the expert in our model provides interpretation for data that originates during the decision-making process. The paper that is closest in terms of how we model expertise is Schwartzstein and Sunderam (2021). They also model the sender's expertise as the knowledge of the distribution of evidence in different states of the world. Their sender persuades a receiver by providing these distributions that the receiver evaluates based on how well they fit the data before using them for decision making. Unlike their model, our auditor evaluates the feasibility of the distribution using its specific consistency properties.

The paper is organized as follows. Section 2 presents the model and states the expert's problem. Section 3 simplifies and restates the expert's problem in terms of justifications. In Section 4, we characterize the expert's strategic justifications. Section 5 studies the impact of the auditor's additional expertise on the expert's strategic justification and the clients' aggregate payoff. We consider an auditor who reasons strategically in Section 6, and conclude with Section 7. Appendices A-D contain proofs of results and lemmas that do not appear in the main text.

2 Model

There is an expert (she), an auditor (he), and a unit mass of clients. Each client has a type (c, ω_c), where $c \in \mathcal{C} = [0, \bar{c}]$ and $\omega_c \in \{l, h\}$. The first element c, observable to both the expert and the auditor, is a payoff relevant parameter that differs across clients. We assume that clients are

[4] See Lehrer and Teper (2011), Cherepanov, Feddersen and Sandroni (2013) for decision theoretic analysis of justifiable preferences.

distributed uniformly on $\mathcal{C}$. The second element ω_c, unknown to all players, is the state of the world associated with the client c. This state is equally likely to be low (l) or high (h) and is drawn independently of the parameter c and states of other clients.[5] For each client c, the expert's task is to diagnose the unknown state ω_c and take action $a_c \in \{L, H\}$, where we refer L as low action.

Payoffs: Payoffs for client c depend on whether the expert's action a_c matches the client's unknown state ω_c. Whenever there is a mismatch, the client gets non-zero payoffs. In particular, if the expert takes action H for a client with $\omega_c = l$ (type-I error), then the client gets a payoff of $-c$. However, when a type-II error happens, the client gets a payoff of -1. The second entry in each cell of the following table gives the payoff of client c.

Action\State	l	h
L	$0, 0$	$-1, -1$
H	$-c + b, -c$	$0, 0$

The expert's payoffs depend on whether her action matches the client's state and whether she can justify her decisions to the auditor. If the expert fails to justify her decisions, she gets a payoff of $-\infty$. Otherwise, her interaction with clients determines her payoffs. When the expert interacts with client c, her payoffs are identical to that of the client, except when the type-I error occurs. In this case, the expert gets a larger payoff $-c + b$, where $b > 0$ captures the conflict of interest. These payoffs are listed as the first entry in each cell of the above table. Unlike the expert and clients, the auditor has no preferences of his own and only performs an audit duty using clients' payoffs.

Information: Both the expert and the auditor know that ω_c is equally likely to be low or high. To estimate this unknown state ω_c associated with client c, the expert obtains a signal (evidence) $y_c \in [0, 1]$. If $\omega_c = l$, then the signal y_c is drawn from a distribution that has full support and a differentiable cumulative distribution function (CDF) F_l such that $F_l^{'}(y_c) \neq 0$. If $\omega_c = h$, the realization y_c is drawn from a distribution with a differentiable CDF F_h. Moreover, the associated density functions f_l, f_h have the Monotone Likelihood Ratio Property (MLRP) and $f_l(y_c)/f_h(y_c) \leq M$.[6] The draw y_c is independent of the draws of other clients.

The expert knows the signal generating process and, therefore, can perform Bayesian inference after observing y_c. On the other hand, the auditor, who also observes y_c, has limited expertise in interpreting it. This is because the auditor does not know the distribution of y_c in the high state of the world. The auditor is only aware that it is some distribution with a differentiable CDF. This

[5] The analysis does not rely on the probabilities of both states being equal. They only have to be non-zero.

[6] For analytical convenience, we assume $M > \bar{c}$.

missing piece from the auditor's knowledge of the signal generating process restricts his ability to interpret commonly observable data. Thus, the expert and the auditor have asymmetric expertise in interpreting commonly observable data.

Relabeling $x_c = F_l(y_c) \in \mathcal{X} = [0,1]$ yields an equivalent and analytically convenient reformulation of the above signal generating process: In the low state, signal x_c is drawn from $U[0,1]$. In the high state, it is drawn from a distribution that has a continuous CDF $F = F_h(F_l^{-1}(x_c))$. Moreover, the associated density $f(x_c) = f_h(F_l^{-1}(x_c))\frac{dF_l^{-1}(t)}{dt}\Big|_{t=x_c}$ is non-decreasing and bounded ($f(x_c) \leq M$).[7] Observe that density function f is also the likelihood ratio function in the current context, and it has MLRP. Intuitively, a higher signal realization x_c is indicative of the higher state of the world. Henceforth, we refer to the function f as *density*.

As before, the expert completely understands the reformulated signal generating process. The auditor knows that x_c is drawn from $U[0,1]$ in the low state. However, in the high state, he only knows that it is drawn from some distribution that has a non-decreasing density J.[8] This density is bounded ($J(x_c) \leq M$) and is atomless (has a continuous CDF). Thus, density J is the part missing from the auditor's knowledge of the signal generating process. Because of such limited knowledge, the auditor only understands the qualitative property that a higher signal realization is indicative of the higher state of the world but does not know the precise extent to which the signal favors one state over the other. Throughout the paper, we impose the restriction that the distribution of x_c in the high state, i.e., F, differs from the uniform distribution.

Expert's decision rule and associated payoffs: The expert makes decision for a client c based on the observed signal x_c. This decision-making can be summarized using a decision rule $d : \mathcal{C} \times \mathcal{X} \to [0,1]$, where $d(c,x)$ denotes the probability of the action H. If the expert follows a decision rule d, then the expected payoff of the expert from client c is

$$\pi_E(c,f,d) = \underbrace{-\frac{1}{2}\int_0^1 (c-b)d(c,x)dx}_{\text{Payoff from action } H} \underbrace{-\frac{1}{2}\int_0^1 (1-d(c,x))f(x)dx}_{\text{Payoff from action } L}. \tag{1}$$

The first integral is the payoff associated with type-I error, and the second integral corresponds to the payoff from type-II error. The total expected payoff of the expert is

$$\Pi_E(f,d) = \frac{1}{\bar{c}}\int_0^{\bar{c}} \pi_E(c,f,d)dc. \tag{2}$$

[7] It is straightforward to verify that $x_c \sim U[0,1]$. The CDF F is continuous because it is a composition of two continuous functions. Moreover, differentiability properties of F_l and F_h ensure that $f(x_c)$ is well-defined. Finally, the equality $f(x_c) = f_h(y_c)/f_l(y_c)$ implies that f is non-decreasing and bounded.

[8] We abuse the convention and refer to a density function with a capital letter.

For client c, the expected payoff $\pi(c, f, d)$ is obtained by equating $b = 0$ in Equation (1). For an arbitrary density function J and a client c, a decision rule that maximizes $\pi(c, J, d)$ is denoted by $d^*_J(c, \cdot)$. Since decision problems of clients are independent of each other, the decision rule d^*_J maximizes the payoff of all clients and is therefore termed as *client-optimal decision rule.*

Timing of interaction and expert's problem: There are three stages to the interaction between the expert, the auditor, and clients.

1. **Choice of a decision rule**: The expert fixes a decision rule $d : \mathcal{C} \times \mathcal{X} \to [0, 1]$.
2. **Expert's decision making**: For each client c, the nature draws the state of the world ω_c. If $\omega_c = l$, then a signal realization x_c is drawn from the distribution $U[0, 1]$. Otherwise, a signal realization x_c is drawn from the distribution F . The expert observe x_c and uses the decision rule d to take the action $a_c(d, x_c) \in \{L, H\}$. Payoffs are accrued to the expert and the client. This decision-making statge gives rise to the data $D = \{x_c, a_c(d, x_c)\}_{c \in \mathcal{C}}$ that is observed by the auditor.
3. **Auditor's scrutiny**: The expert provides a justification $J : \mathcal{X} \to \mathbb{R}$ for the data $D = \{x_c, a_c(d, x_c)\}_{c \in \mathcal{C}}$, and the auditor verifies if data D is justifiable by the reported justification J. If data D is justifiable by J, the interaction ends with the expert getting the payoff accrued in the previous stage of the interaction. Otherwise, the expert gets a payoff of $-\infty$. To complete the description of the interaction, consider the following definition.

Definition 1 (Justifiable Data)**.** *Data D is justifiable by a justification $J : \mathcal{X} \to \mathbb{R}$ if*

a) ***Feasibility****: J is a non-decreasing density function that is bounded and atomless,*

b) ***Consistency with clients' interests****: there exists a client-optimal decision rule d^*_J such that for every $x_c \in D$, the expert's action $a_c(d, x_c)$ belongs to the support of $d^*_J(c, x_c)$.*

Thus, a justification J is the expert's claim about the distribution of signal x_c conditional on the state of the world being high. The auditor is boundedly rational, relying on the expert's justification to learn the relevant distribution. First, the auditor verifies if this justification is consistent with his knowledge about f. A justification that passes this first round of scrutiny is termed a *feasible justification.* Then, the auditor uses the signal generating process obtained using justification J to verify if the expert's behavior in data is consistent with the maximization of clients' interests.

The auditor's reliance on the expert to fill the gap in his knowledge may be due to his lack of ability or resources to obtain the necessary knowledge. Moreover, this reliance may also result from rules and protocols that prevent the auditor from using her personal beliefs during scrutiny and restrict him to use only the objective data and the expert's claim.

The expert wants to choose a decision rule in the first stage of the interaction that maximizes her payoff in the second state while ensuring that the expert always passes the auditor's scrutiny. To avoid the severe penalty associated with failing the expert's scrutiny the expert must choose a decision rule that always gives rise to a justifiable data D.

Definition 2 (Justifiable Decision Rule)**.** *A decision rule d is justifiable if each data D originating from it is justifiable by some justification J_D.*

Thus, the expert's problem is

$$\begin{aligned} \max_{d} \quad & \Pi_E(f, d) \\ \text{subject to} \quad & d \ \text{ is justifiable.} \end{aligned} \tag{3}$$

The set of justifiable decision rules is non-empty as it contains d^*_f: the expert always follows the client-optimal decision rule with respect to the actual density f and reports it truthfully as a justification.

In the remaining paper, we ask if and when the restriction of using only justifiable decision rules forces the self-interested expert to act in the client's best interest.

3 Simplifying expert's problem

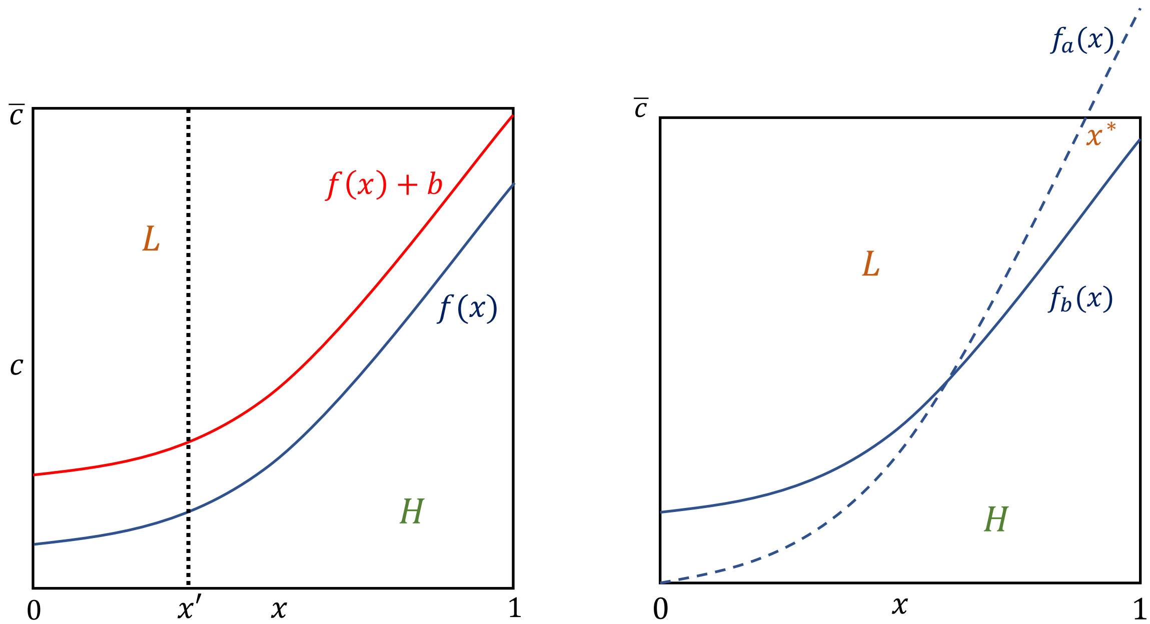
This section characterizes the expert's behavior permissible under a justifiable decision rule. For this purpose, we examine two benchmark decision rules associated with the actual density f: the client-optimal decision rule d^*_f, and the expert's unconstrained optimal decision rule d^E_f. This section establishes that these rules are threshold rules and uses their threshold nature to simplify the expert's problem.

Let density J be any non-decreasing, bounded density that has a continuous CDF. Henceforth, we refer to it either as a *feasible density* or as a *feasible justification.* The following observation obtains the two benchmark decision rules for any feasible density J.

Observation 1. *Fix a feasible density J.*

a) The expert's unconstrained optimal decision rule is a threshold rule

$$d^E_J(c, x) = \begin{cases} 1 & c \leq J(x) + b \\ 0 & c > J(x) + b \end{cases}.$$



Panel (a): Benchmark decision rules   Panel (b): Value of knowledge

Figure 2: Graphical illustration of benchmark decision rules.

b) The client-optimal decision rule is a threshold rule

$$d_J^*(c,x) = \begin{cases} 1 & c \leq J(x) \\ 0 & c > J(x) \end{cases}.$$

Proof. Observe that the total payoff of the expert (Equation (2) for the valid density J) is an uniform aggregation of $\pi_E(c, J, d)$ across all clients. Thus, it suffices to find a decision rule that maximizes $\pi_E(c, J, d)$, where

$$\pi_E(c, J, d) = \int_0^1 \left[-\frac{1}{2}(c-b)d(c,x) - \frac{1}{2}J(x)(1-d(c,x)) \right] dx.$$

For a fixed signal realization x, the integrand in the above expression is a convex combination of $-\frac{1}{2}(c-b)$ and $-\frac{1}{2}J(x)$. The decision rule d_J^E puts all the weight on the larger term and resolves indifference in favor of action H. Since the decision rule d_J^E point-wise maximizes the integrand, it also maximizes $\pi_E(c, J, d)$. Part (b) is obtained by letting $b = 0$ in part (a). □

Panel (a) in Figure 2 pictorially depicts the expert's unconstrained optimal decision rule d_f^E and the client-optimal decision rule d_f^* for the actual density f. For any given signal realization x, the expert has a threshold client $f(x) + b$ such that the expert prefers action H for all clients with payoff parameters smaller than the threshold. Otherwise, the expert prefers action L. Similarly, the auditor prefers action H for clients weakly below $f(x)$.

Any decision rule that differs from d_J^* on a set of signal realizations with zero Lebesgue measure is also client-optimal. However, to maintain the tractability, we assume that the auditor considers d_J^* as the only client-optimal decision rule, and the expert is aware of the auditor's considerations. This focus on deterministic threshold rules has three important implications.

First, the unconstrained expert prefers to take action H for a weakly larger set of clients than the auditor for any signal realization x. For example, the zone of conflict between the expert and the auditor is given by the area between two curves, $f(x)$ and $f(x) + b$, in Panel (a) of Figure 2.

To see the second implication of the threshold rule, consider the following definition. The expert's knowledge is *valuable for a signal realization* x if there exists a client c such that $a_c(d_f^*, x) = L$. For example, in Panel (b) of Figure 2, the expert's knowledge of f_a is valuable at all signal realizations $x < x^*$. However, if $x \geq x^*$, then the high action is client-optimal for all clients $c \in \mathcal{C}$. In other words, the expert's discriminating knowledge is of no use to clients for these latter signal realizations. Using this definition of a valuable signal realization, we define the following property that will eventually characterize the expert's optimal behavior.

Definition 3 (Always Valuable Knowledge)**.** *The expert's knowledge is always valuable on* $[\underline{x}, \overline{x}] \subset [0, 1]$ *if it is valuable for signal realizations* $[\underline{x}, \overline{x})$.

For the density f_b, shown in Panel (b) of Figure 2, the expert's knowledge is always valuable on $[0, 1]$. For any subset $[\underline{x}, \overline{x}] \subset [0, 1]$, the density f is such that either the expert's knowledge is always valuable on $[\underline{x}, \overline{x}]$, or it is not. In section 4, we will show that this classification will characterize the expert's behavior.

To see the third and final implication of the deterministic threshold rule, fix a justifiable decision rule d. And, define data $D(x') = \{x_c = x', a_c(d, x')\}_{c\in\mathcal{C}}$ using a signal realization x'. The dotted line in Panel (a) of Figure 2 graphically represents the first argument of this data. Since data $D(x')$ originates from a justifiable decision rule d, it is justifiable with some feasible justification $J_{D(x')}$ that makes actions $\{a_c(d, x')\}_{c\in\mathcal{C}}$ appear as client-optimal. We know from part (b) of Observation 1 that in any client-optimal decision rule the threshold-client is given by $J(x)$. Thus, we must have $a_c(d, x') = H$ for all $c \leq J_{D(x')}$ and $a_c(d, x') = L$ for all $c > J_{D(x')}$. In other words, the decision rule d must be deterministic threshold rule as stated in the following observation.

Observation 2. *Fix a justifiable decision rule d and a signal realization $x' \in [0,1]$. Then,*

$$a_c(d, x') = \begin{cases} H & c \leq J_{D(x')}(x') \\ L & c > J_{D(x')}(x') \end{cases}.$$

To see a deeper implication of the above observation, consider the function $\hat{J}(x) = J_{D(x)}(x)$, obtained by point-by-point piecing feasible densities $\{J_{D(x)}\}_{x \in [0,1]}$. For a given signal realization x, comparing the threshold client $J_{D(x)}(x)$ in Observation 2 to the threshold client of d^*_J (part (b) of Observation 1) suggests that a justifiable decision rule d can be thought of as a client-optimal decision rule for the function $\hat{J}(x)$. Thus, if the function $\hat{J}$ turns out to be a feasible density function, then the expert using the decision rule d can provide $\hat{J}$ as a justification for all realizations of the data. Equivalently, the decision rule d satisfies the following stronger notion of justifiability.

Definition 4 (Ex-ante Justifiable Decision Rule)**.** *A decision rule d is ex-ante justifiable if every data D originating from it is justifiable by the same justification J.*

To understand why this is a stronger notion, recall that the definition of a justifiable decision rule d allowed the expert to report different feasible densities (justifications) for different data realizations. The ex-ante justifiability, on the contrary, means that the expert can report the same feasible density J as a justification for all possible realizations of the data. The following theorem asserts that the weaker notion of justifiability implies ex-ante justifiability.

Theorem 1. *If a decision rule d is justifiable then it is ex-ante justifiable.*

The proof of the above theorem involves showing that the function $\hat{J}$ indeed a feasible justification and is relegated to the Appendix C. The above theorem captures the restrictions imposed on the expert's behavior by a justifiable decision rule. Intuitively, the use of justifiable decision rules forces the expert to follow the client-optimal decision rule d^*_J, albeit with any feasible density J and not necessarily f. Thus, we can restate the expert's problem (3) as

$$\begin{aligned} &\max_J \ \frac{1}{\bar{c}} \int_0^{\bar{c}} \pi_E(c, f, d^*_J) dc \\ \text{subject to } \ &J \ \text{ is a non-decreasing, bounded, and atomless function} \\ &\int_0^1 J(x) dx = 1. \end{aligned} \tag{4}$$

Using the definition of the client-optimal decision rule d^*_J stated in Observation 1, we can further simplify the expert's problem.

Observation 3. *The expert's optimization problem can be written as*

$$\max_J \; -\int_0^1 \left[J_{\overline{c}}(x) \left(\frac{J_{\overline{c}}(x)}{2} - b \right) + (\overline{c} - J_{\overline{c}}(x)) f(x) \right] dx \tag{P}$$

$$\textit{subject to } \; J \; \textit{ is a non-decreasing, bounded, and atomless function}, \qquad \text{(monotonicity)}$$

$$\int_0^1 J(x)dx = 1, \qquad \text{(aggregate consistency)}$$

where $J_{\overline{c}}(x) = min\{J(x), \overline{c}\}$.

Section 4 discusses the solution to the above problem, and we refer to it as the expert's *strategic justification*. Before moving towards identifying the strategic justification, we examine cases of few feasible and infeasible justifications. Observe that the justification $J = f$ satisfies both the monotonicity and aggregate consistency constraint and thus is feasible. Justification $J(x) = f(x)+b$ satisfies monotonicity but violates aggregate consistency. In fact, aggregate consistency precludes justifications obtained by inflating f on a positive Lebesgue measure of signal realizations. Thus, if a strategic justification lies above f at some signal realizations, then it must lie below f for some other signal realizations.

4 Strategic justifications

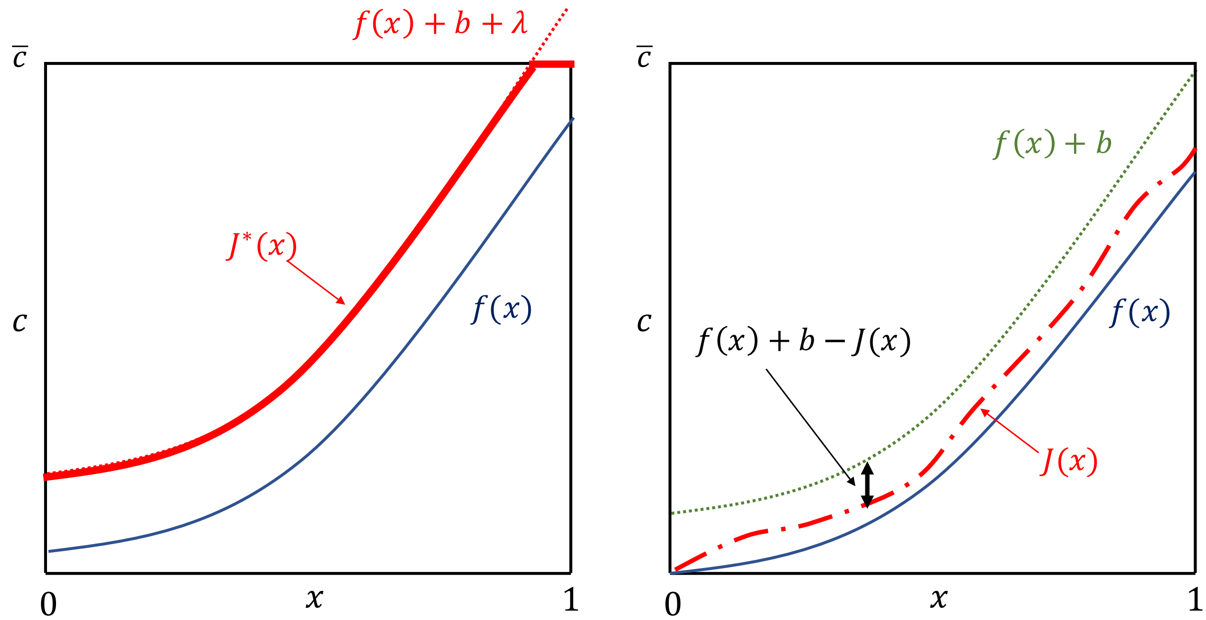
This section analyzes the expert's strategic justifications. It identifies necessary and sufficient conditions for the strategic justification to equal the actual density f almost everywhere. Consequently, it answers if and when the constraint of using justifiable decision rules forces the expert to act in the clients' best interest.

To obtain the strategic justification, we start by generalizing the problem in Observation 3 in two aspects. First, we focus attention on any sub-interval $[\underline{x}, \overline{x}] \subset [0, 1]$ of the signal realizations. Second, the expert's justification is allowed to integrate to m, where $\int_{\underline{x}}^{\overline{x}} f(x)dx \leq m \leq \int_{\underline{x}}^{\overline{x}} \min\{f(x)+b, \overline{c}\}dx$. These bounds eliminate values of m that are irrelevant for an optimizing expert. The following theorem characterizes the shape of the strategic justification on the interval $[\underline{x}, \overline{x}]$.

Theorem 2. *Suppose that the expert's knowledge is always valuable on* $[\underline{x}, \overline{x}]$*. Then the expert's strategic justification* $J^*(x) : [\underline{x}, \overline{x}] \to \mathbb{R}_+$ *is given by*

$$J^*(x) = min\{f(x) + b + \lambda, \overline{c}\}, \tag{5}$$

where $-b \leq \lambda \leq 0$ *and* $\int_{\underline{x}}^{\overline{x}} J^*(x)dx = m$.



Panel (a): Strategic Justification

Panel (b): Non-optimal J

Figure 3: Strategic justification when the expert's knowledge is always valuable on $[0, 1]$.

The thick curve (red) in Panel (a) of Figure 3, pictorially illustrates the form of strategic justification $J^*(x)$ for $\underline{x} = 0$, $\overline{x} = 1$ and appropriate $1 < m < 1 + b$. The strategic justification $J^*(x)$ is obtained by projecting the function $f(x) + b + \lambda$ on the area below level $\overline{c}$, where the constant λ is chosen such that the aggregate consistency constraint is satisfied. Henceforth, $J^m_S(x)$ denotes the strategic justification obtained by applying Theorem 2 to interval $S \subset [0, 1]$ with mass m.

To understand why translation λ is a constant and not some arbitrary function of signal realization x, consider the case $\underline{x} = 0$, $\overline{x} = 1$ and $f(1) + b < \overline{c}$. Panel (b) of Figure 3 illustrates these conditions graphically. Moreover, $1 < m < 1 + b$. In this case, the expert has a mass $m - 1$ that can be used to inflate the justification above the density f at various signal realizations. For any signal realization x, inflating the strategic justification above the density f allows the expert to replace action L (type-II error) with more profitable action H (type-I error). The marginal benefit of such switch can be shown to equal the difference $f(x) + b - J(x)$. In the strategic justification the expert equates this marginal benefit across all signal realizations and spreads the mass $m - 1$ evenly. Thus, $\lambda = m - 1 - b < 0$

In the above intuition, if $m = 1$, then the expert has no additional mass to inflate the strategic justification above the actual density f. Thus, the strategic justification collapses to the actual density f, and the expert acts in the best interest of all clients. Thus,

Proposition 1. *If the expert's knowledge is always valuable on* $[0,1]$*, then the expert's strategic justification coincides with* f *almost everywhere on* $[0,1]$*.*

Proof. Since the expert's knowledge is always valuable on $[0,1]$, we have $f(x) < \bar{c}$ for $x \in [0,1)$. Applying Theorem 2 for the interval $[0,1]$ and $m = 1$, we obtain

$$J^*(x) = \min\{f(x) + b + \lambda, \bar{c}\},$$

where $-b \leq \lambda \leq 0$ and $\int_0^1 J^*(x)dx = 1$. To complete the proof, observe that any $\lambda \neq -b$ violates the aggregate consistency. □

When the expert's knowledge is always valuable, the expert wants to inflate the justification above the actual density f at all signal realizations. However, such a justification violates the aggregate consistency constraint. Thus, if a strategic justification lies above f on a non-zero measure of signal realizations, then it must lie below f for a non-zero measure of signal realizations. As the discussion following Theorem 2 suggests, such a justification obtained by simultaneously inflating and deflating density f is not optimal because it does not equate the marginal benefit $f(x)+b-J(x)$ across all signal realizations. Therefore, the strategic justification equals the actual density f almost everywhere, and the expert acts in the clients' best interest. In other words, a link is created across various decisions (identified by the pair (c, x_c)) of the expert by the consistency requirements of a feasible justification. This link is "perfect" when the expert's knowledge is always valuable on $[0,1]$.

The mechanism behind Proposition 1 is reminiscent of the "quota" mechanisms (Jackson and Sonnenschein (2007)). In the context of asymmetrically observed data, these mechanisms link the expert's incentives across many independent copies of the same decision problem. However, they rely on the auditor perfectly knowing the entire generating process of the expert's privately observed data and therefore are not robust. Proposition 1 identifies sufficient conditions for a similar mechanism to work in the context of symmetrically observed data and asymmetric expertise. This mechanism works even when only specific properties of the data generating process rather than the complete process are common knowledge among the players.

We now show that if the expert's knowledge is not always valuable on $[0,1]$ and f is "generic", the link created by justifiable decision rules is "imperfect." The expert can further own interest while maintaining justifiability.

Suppose that the expert's knowledge is not always valuable on $[0,1]$. Let $V = \{x : f(x) < \bar{c}\}$, denote the set of signal realizations where the expert's knowledge is valuable. Since f is non-decreasing, the set V is given by either $V = [0, x^*)$ or $V = [0, x^*]$, where $x^* < 1$. For brevity, we henceforth assume that $V = [0, x^*)$.[9] We say f is generic if it satisfies the following assumption.

Assumption 1. *f is such that $x^* > 0$ and f differs from $\bar{c}$ on V^c.*[10]

The density function f_a in panel (b) of Figure 2 satisfies the above assumption. If $x^* = 0$, then the first part of Assumption 1 is violated and set V is empty. In this case, $f(x) \geq \bar{c}$ for every $x \in [0,1]$. Thus, for any signal realization x, both the expert and the auditor want the same action H for all clients in set $\mathcal{C}$. In other words, there is no effective conflict between the expert and clients, and the expert always acts in the clients' best interest.

When f satisfies the first part of the Assumption 1, both V and V^c are non-empty intervals. For a signal realization $x \in V^c$, the expert does not benefit by inflating the justification above the actual density f. This is because there are no clients with parameter $c > f(x)$ for whom the expert can take her preferred action. In fact, the expert can deflate the justification up to the level $\bar{c}$ and still continue to take her preferred action H for all clients. By performing such deflation for all signal realizations in set V^c, the expert can create a *supply* of the misrepresentation

$$S(f) = \int_{x^*}^{1} (f(x) - \bar{c})\, dx \geq 0.$$

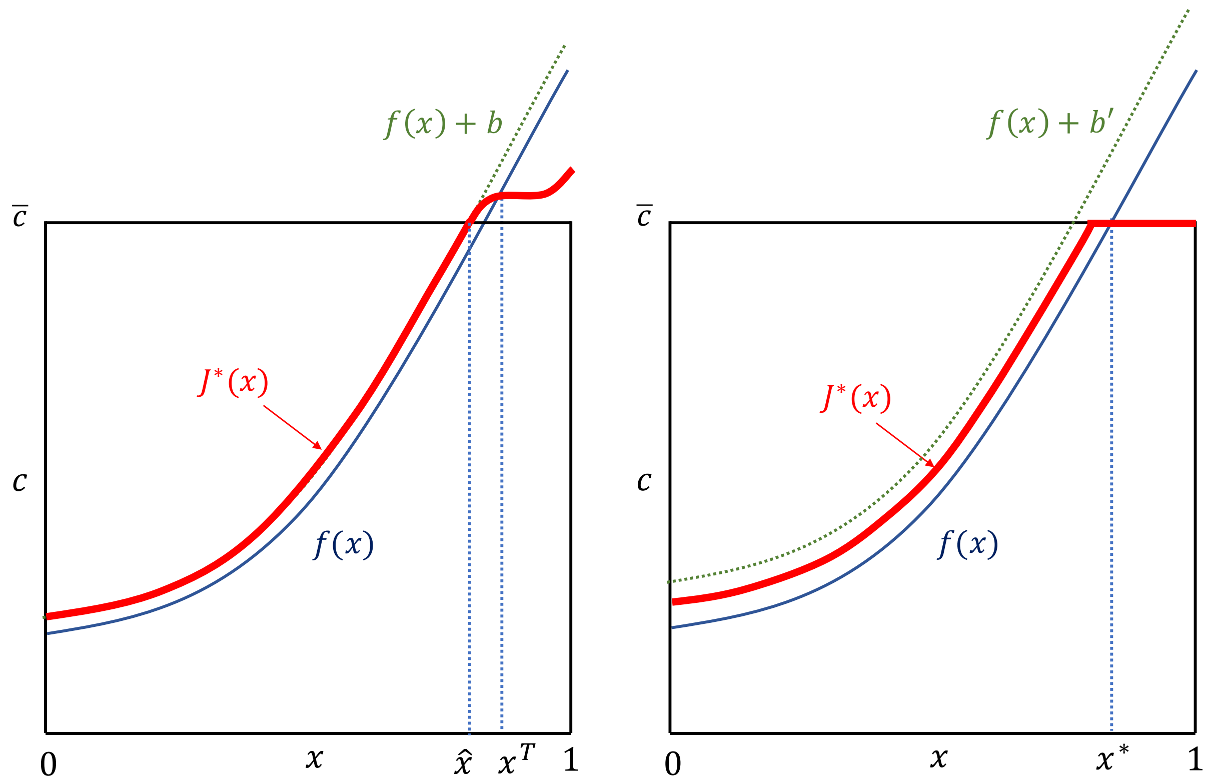
The second part of Assumption 1 guarantees that supply $S(f)$ is non-zero. The expert can possibly use some of the mass $S(f)$ to inflate the justification above the actual density f for signal realizations in set V. Such operation increases the expert's payoff as it allows her to take her preferred action H on larger mass of (c, x_c) pairs. Thus, the expert's *demand* of misrepresentation on these signals is given by

$$D(f) = \int_{0}^{x^*} (\min\{f(x) + b, \bar{c}\} - f(x))\, dx \geq 0.$$

The first part of Assumption 1 ensures that the Demand $D(f)$ is non-zero. The following proposition shows that the expert's strategic justification J^* transfers mass $\Delta = \min\{D(f), S(f)\}$ from signal realizations V^c to signal realizations V. And, spreads the mass $m = \int_V f(x)dx + \Delta$ on signal realizations V using the optimal justification J_V^m obtained through Theorem 2.

[9] All the results and their proofs readily extend to the case $V = [0, x^*]$.

[10] Function h differs from function g on set $A \subset [0,1]$ if they differ on a subset of A that has a positive Lebesgue measure.



Panel (a): $D(f) \leq S(f)$ Panel (b): $D(f) > S(f)$

Figure 4: Strategic justification when the expert's knowledge is not always valuable on $[0, 1]$.

Proposition 2. *Suppose the expert's knowledge is not always valuable on $[0, 1]$ and f is generic. Then there exists $x_T \geq x^*$, such that the expert's strategic justification J^* lies above the actual density f for signal realizations smaller than x_T and otherwise lies weakly below f. In particular,*
(a) If $D(f) \leq S(f)$, then $J^*(x) = f(x) + b$ for x such that $f(x) + b < \bar{c}$. Moreover, $x_T \geq x^*$.
(b) If $D(f) > S(f)$, then $x_T = x^*$ and

$$J^*(x) = \begin{cases} J_V^m(x) & x < x^* \\ \bar{c} & x \geq x^*, \end{cases}$$

where $m = \int_V f(x)dx + \Delta$.

The thick (red) curve in Panel (a) of Figure 4 exhibits the strategic justification J^* when the expert (with $b > 0$) is demand-constrained ($D(f) \leq S(f)$). The part of the strategic justification corresponding to signal realizations larger than $\hat{x}$ can be any non-decreasing function that lies above $\bar{c}$ and respects the aggregate consistency constraint. If the expert (with $b' > b$) is supply-constrained, then the thick (red) curve in Panel (b) of Figure 4 exhibits the associated strategic justification.

In both these strategic justifications, the expert inflates the strength of the evidence weaker than x_T, and deflates the strength of the evidence stronger than x_T. Such a misrepresentation harms a client c with borderline signal realizations $x_c \in [f^{-1}(c-b), f^{-1}(c))$.

Proof of Proposition 2. Suppose $D(f) \leq S(f)$. For any signal realization $x \in V^c$, the expert wants to take action H for all clients in set $\mathcal{C}$. Doing so in a justifiable manner, however, requires only a mass $\bar{c}(1-x^*)$. Since this mass is strictly smaller than $\int_{x^*}^1 f(x)dx$ (Assumption 1), there exists a supply $S(f) > 0$ of mass. On the other hand, taking the unconstrained optimal action on signal realizations in the set V in a justifiable manner requires a mass of $\int_0^{x^*} f(x)dx + D(f)$. Since $D(f) \leq S(f)$, the expert can transfer mass $D(f)$ to signal realizations in set V and take her unconstrained optimal actions for every $(c, x_c) \in \mathcal{C} \times \mathcal{X}$. In particular, $J^*(x) = f(x) + b$ for $x < \hat{x}$, where $\hat{x}$ is the smallest signal realization such that $f(x) + b = \bar{c}$. For signal realizations $x \geq \hat{x}$, the strategic justification J^* distributes the remaining mass $1 - \int_0^{\hat{x}}(f(x)+b)dx$ using any non-decreasing, bounded and atomless function $g(x)$ that satisfies the following conditions. First, $g(x) \geq \bar{c}$ for $x \geq \hat{x}$. Second, the resulting J^* must respect the aggregate consistency constraint. The threshold signal realization $x_T \geq x^*$ depends on the exact details of $g(x)$.

Now consider the case, $D(f) > S(f)$. Let J be a feasible justification such that $J(x) > \bar{c}$ on a non-empty interval $I \subset [0,1]$. Since the expert is supply-constrained, there exist a positive mass of signal realizations in V where $J(x) < \min\{f(x)+b, \bar{c}\}$. Reducing $J(x)$ to $\bar{c}$ on I and using the mass $\int_I J(x)dx - \bar{c}|I|$ to inflate the justification on aforementioned signal realizations in V increases the expert's payoff. Therefore, J is dominated. This observation implies that the strategic justification must satisfy $J^*(x) \leq \bar{c}$ for all $x \in [0,1)$. Using Lemma 5, the strategic justification has the form $J^*(x) = \min\{f(x)+b+\lambda, \bar{c}\}$, where $-b \leq \lambda \leq 0$ and $\int_0^1 J^*(x)dx = 1$.

We must have $-b < \lambda < 0$. The left inequality holds because for $\lambda = -b$ the aggregate consistency constraint is violated. To see why, notice that the inequality $J^*(x) \leq f(x)$ holds for all $x \in [0,1]$, and holds strictly for $x \in V^c$. Similarly, $\lambda = 0$ contradicts that the expert is supply-constrained. We complete the proof by showing that the strategic justification J^* matches the form proposed in the statement of the proposition, when $-b < \lambda < 0$.

Since $f(x) \geq \bar{c}$ for $x \geq x^*$, and $\lambda + b > 0$, we have $f(x) + b + \lambda > \bar{c}$. Thus, $J^*(x) = \bar{c}$ on $[x^*, 1]$. With strategic justification $J^*(x)$, the expert frees up a mass of $S(f)$ by deflating the density f on $[x^*, 1]$. The strategic justification J^* must spread the mass $m = \int_V f(x)dx + S(f)$ in an optimal manner on V. By Theorem 2, the strategic justification that spreads of mass m on $V = [0, x^*)$ in an optimal manner is given by $J_V^m(x)$. Thus, $J^*(x) = J_V^m(x)$ on $[0, x^*)$. □

We conclude our discussion of the expert's strategic justifications by answering the following question. When does the restriction of using only justifiable decision rules force the self-interested expert to take the client-optimal action for almost all clients-signal-realization pairs? Putting Propositions 1 and 2 together, we obtain

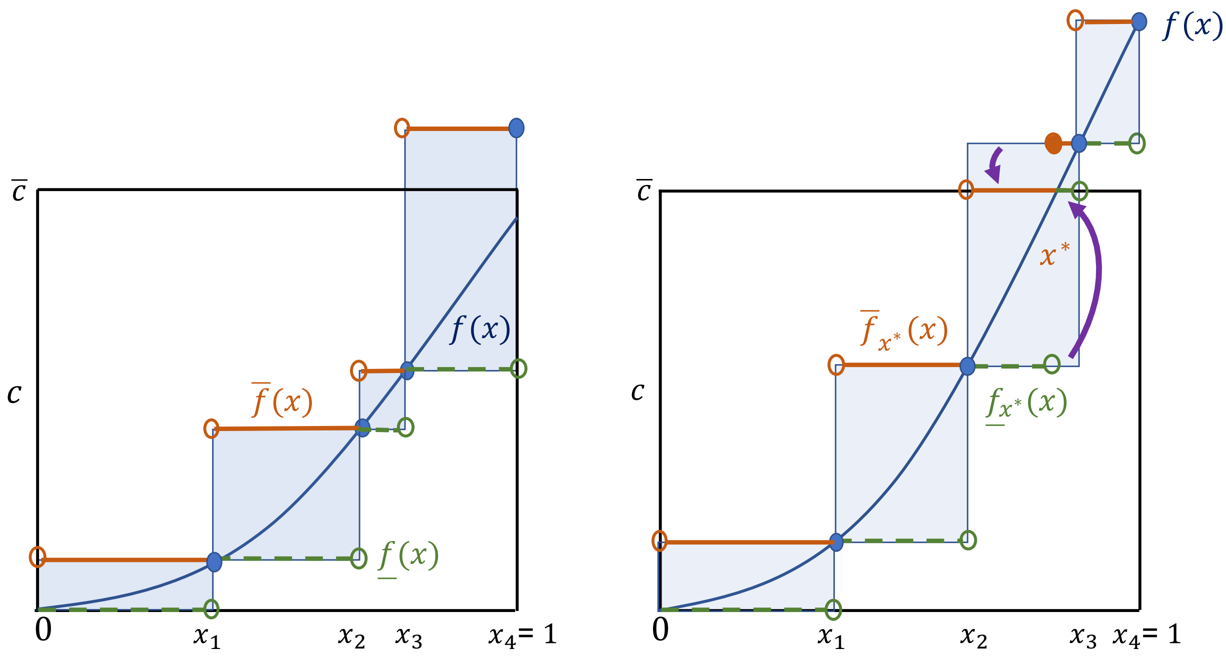
Corollary 1. *Suppose f is generic. Then the expert takes the client-optimal action for almost all clients-signal-realization pairs iff the expert's knowledge is always valuable on* $[0, 1]$.

Proof. To complete the proof, it suffices to show that f satisfies Assumption 1 if the expert's knowledge is always valuable on $[0, 1]$. Suppose that the expert's knowledge is always valuable on $[0, 1]$. Then, we either have $V = [0, 1)$ or $V = [0, 1]$. In the first case, V^c has zero Lebesgue measure. Moreover, V^c is empty in the latter case. Thus, f is generic. □

5 Value of auditor's expertise

So far, we considered an auditor who deemed any justification J as a feasible justification as long as it satisfied specific consistency properties (Part (a) in Definition 1). However, the auditor may know more about the distribution of the signal in the high state of the world than just a few consistency properties. For example, the surgeon on the medical board may have had experience interpreting certain symptoms and test results. A key feature of such additional knowledge usually obtained through case studies and past experiences is that it makes the auditor more confident in interpreting certain signal realizations than others. To model the auditor's expertise that is variable across signal realizations, we equip the auditor with a set K of *anecdotes.* Formally, $K = \{x_1, x_2, \cdots, x_k\}$, where k is finite and $x_k = 1$. Anecdotes are signal realizations where the auditor knows the value of f precisely, and therefore, has the same expertise as the expert in interpreting these signal realizations. However, for intermediate signal realizations, the auditor can only infer f imperfectly. We assume that the expert knows the auditor's set of anecdotes K.

In this section, we first identify the expert's strategic justifications when the auditor has anecdotes. Section 5.1 shows that the key results of the previous section extend to this setting with anecdotes. We then explore whether increasing the expertise of the auditor through more anecdotes benefits clients'. The auditor's anecdotal expertise has no benefit when the expert's knowledge is always valuable on $[0, 1]$. This is unsurprising as the expert's strategic justification coincided with the actual density f even without anecdotes. However, the analysis is more nuanced when the expert's



Panel (a): Always valuable knowledge

Panel (b): Knowledge of limited value

Figure 5: Graphical representation of bounds $\underline{f}$, $\overline{f}$, and modified bounds $\overline{f}_{x^*}$ and $\underline{f}_{x^*}$.

knowledge is not always valuable on $[0, 1]$. Section 5.2 identifies conditions on the configuration of the additional anecdotes such that they reduce clients' aggregate payoff.

5.1 Strategic justifications with anecdotes

An auditor with anecdotes in set K requires a *feasible justification*, in addition to the requirements of part (a) of Definition 1, to satisfy $J(x) = f(x)$ for $x \in K$. This additional restriction strengthens the definition of *justifiable data*. To see how anecdotes restrict the set of feasible justifications, consider the inference conducted by the auditor who combines anecdotes with the fact that f is a non-decreasing function. Such auditor infers that f must fall between two functions, $\underline{f}(x)$ and $\overline{f}(x)$. Panel (a) in Figure 5 pictorially depicts $\underline{f}$ and $\overline{f}$ for $K = \{x_1, x_2, x_3, x_4 = 1\}$. The dashed (green) step function is $\underline{f}(x)$, whereas the solid (brown) step function illustrates $\overline{f}(x)$. The formal definitions of these functions depend on the set of anecdotes K and are given in Appendix A. Observe that both these functions, $\underline{f}$ and $\overline{f}$, coincide at $x \in K$. Even in the presence of these anecdotes, we can write the expert's problem in terms of justification.

Observation 4. *The expert's optimization problem is obtained by adding the following constraint* *to the problem (P) stated in Observation 3.*

$$\underline{f}(x) \leq J(x) \leq \overline{f}(x) \quad \text{for } x \in [0, 1]. \qquad \text{(local consistency)}$$

The constraint in Observation 4 is termed local consistency because it captures each anecdote's local effect in combination with the monotonicity property of feasible justification. This constraint severely restricts the expert's ability to inflate the strategic justification at the anecdote and at signal realizations just below it. The affine nature of local consistency allows us to extend Theorem 2 to the setting with anecdotes. The following theorem gives the expert's strategic justification on the interval $[\underline{x}, \overline{x}]$ for $\int_{\underline{x}}^{\overline{x}} f(x)dx \leq m \leq \int_{\underline{x}}^{\overline{x}} \min\{f(x) + b, \overline{f}(x), \overline{c}\}dx$.

Theorem 2A. *Suppose that the expert's knowledge is always valuable on $[\underline{x}, \overline{x}]$. Then the expert's strategic justification $J^*(x) : [\underline{x}, \overline{x}] \to \mathbb{R}_+$ is given by*

$$J^*(x) = min\left\{f(x) + b + \lambda, \overline{f}(x), \overline{c}\right\}, \tag{6}$$

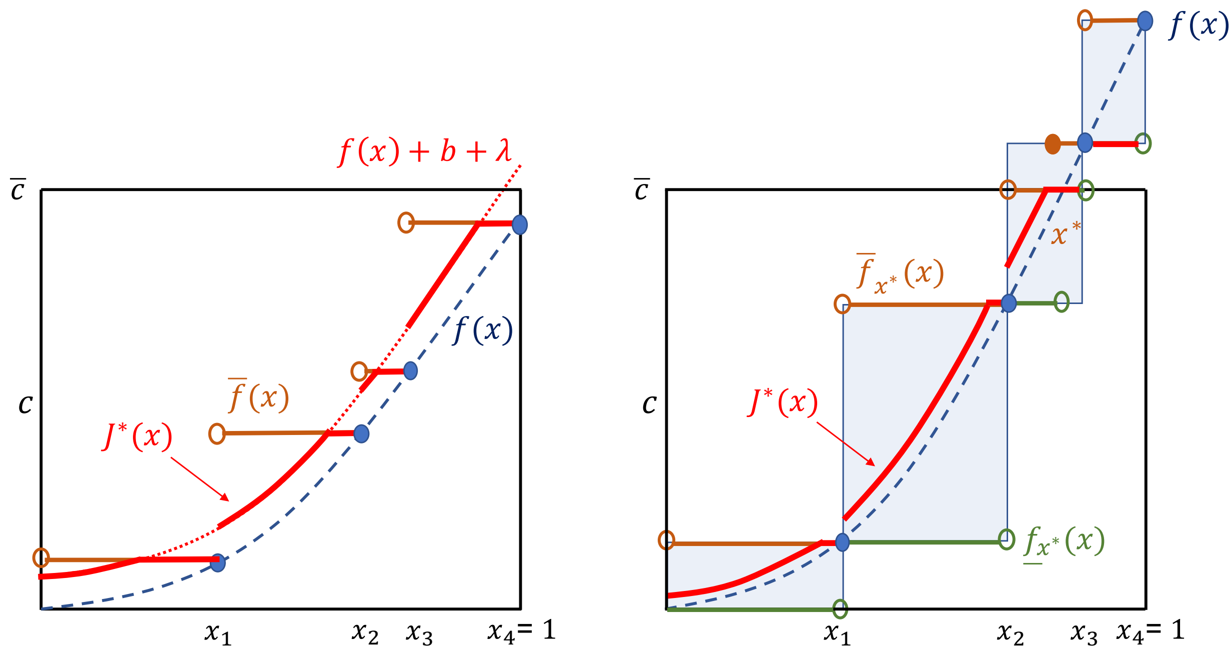
where $-b \leq \lambda \leq 0$ and $\int_{\underline{x}}^{\overline{x}} J^*(x)dx = m$.

The solid (red) curve in Panel (a) of Figure 6, pictorially illustrates the form of strategic justification $J^*(x)$ for $\underline{x} = 0$, $\overline{x} = 1$ and appropriate $m > 1$. The strategic justification $J^*(x)$ is obtained using projecting the function $f(x) + b + \lambda$ on the area below the curve $\min\{\overline{f}(x), \overline{c}\}$. To see how anecdotes change the strategic justification, compare solid (red) curves in Panel (a) of Figure 3 and Figure 6. The direct effect of the anecdotes brings the strategic justification closer to the actual density f at the anecdote and at signal realizations just below it. However, to maintain the aggregate consistency, the strategic justification may move away from the actual density f at signal realizations away from the anecdotes. Thus, the net effect is given by the combination of the direct and the indirect effect.

The interplay of both these effects is most visible when the expert's knowledge is not always valuable on $[0, 1]$. This is because, if the expert's knowledge is always valuable on $[0, 1]$, then the strategic justification coincides to the actual density f even without anecdotes. Thus, Proposition 1 continues to hold even with anecdotes. To study the interplay of the direct and the indirect effect of anecdotes, the remaining section considers the case when the expert's knowledge is not always valuable on $[0, 1]$.

If the expert's knowledge is not always valuable on $[0, 1]$, then there exists $x^* < 1$ such that $V = [0, x^*)$[11]. We modify bounds $\underline{f}$ and $\overline{f}$ to obtain a new lower bound $\underline{f}_{x^*}(x)$ and an upper bound

[11] As before, we omit the case where $V = [0, x^*]$ for the brevity.



Panel (a): Always valuable knowledge

Panel (b): Knowledge of limited value

Figure 6: Shape of strategic justifications with anecdotes

$\overline{f}_{x^*}$ in the following manner. Let $x^* \in (x_{j-1}, x_j]$, where $j \leq k$ and $x_{j-1}, x_j \in K$. If $j = 1$ then $x_{j-1} = 0$. Define

$$\underline{f}_{x^*}(x) = \begin{cases} \max\{\bar{c}, \underline{f}(x)\} & V^c \cap (x_j, x_{j+1}] \\ \underline{f}(x) & \text{otherwise} \end{cases} \qquad \overline{f}_{x^*}(x) = \begin{cases} \min\{\bar{c}, \overline{f}(x)\} & x \in V \cap (x_j, x_{j+1}] \\ \overline{f}(x) & \text{otherwise} \end{cases}.$$

Panel (b) in Figure 5 graphically illustrates $\underline{f}_{x^*}(x)$ and $\overline{f}_{x^*}$. In this case, $j = 3$ and $x^* \in (x_2, x_3]$. Observe that the definition of $\underline{f}_{x^*}$ differs from $\underline{f}$ only on $(x^*, x_3]$. Similarly, definition of $\overline{f}_{x^*}$ differs from $\overline{f}$ only on $(x_2, x^*]$. To obtain the expert's strategic justification in this case, assume that f satisfies the following assumption (letter A stands for anecdotes) analogous to Assumption 1.

Assumption 1A. *f is such that $x^* > 0$. Moreover, f differs from $\overline{f}_{x^*}(x)$ on $[0, x^*)$ and from $\underline{f}_{x^*}(x)$ on $[x^*, 1]$.*

The density f in the Panel (b) of Figure 5 satisfies the above assumption. Any f that satisfies Assumption 1A has the $D(K, f), S(K, f) > 0$, where

$$D(K, f) = \int_0^{x^*} \left(\min\{f(x) + b, \overline{f}_{x^*}(x)\} - f(x)\right) dx, \qquad S(K, f) = \int_{x^*}^1 \left(f(x) - \underline{f}_{x^*}(x)\right) dx.$$

As before, the expert does not benefit by inflating the justification above the actual density f on signal realizations $x \geq x^*$. Thus, the expert can transfer mass $\Delta = \min\{D(K,f), S(K,f)\}$ from signal realizations V^c to signal realizations V. And can spread the mass $m = \int_V f(x)dx + \Delta$ optimally on signal realizations V using the justification J_V^m (obtained by applying Theorem 2A to set V with mass m). The following proposition shows that such a heuristic yields the expert's strategic justification.

Proposition 2A. *Suppose the expert's knowledge is not always valuable on* $[0,1]$ *and* f *satisfies Assumption 1A. Then there exists* $x_T \geq x^*$*, such that the expert's strategic justification* J^* *lies above the actual density* f *for signal realizations smaller than* x_T *and otherwise lies weakly below* f*. In particular,*

(a) If $D(K,f) \leq S(K,f)$*, then* $J^*(x) = \min\{f(x)+b, \overline{f}_{x^*}(x)\}$ *for* x *such that* $\min\{f(x)+b, \overline{f}_{x^*}(x)\} < \bar{c}$*. Moreover,* $x_T \geq x^*$*.*

(b) If $D(K,f) > S(K,f)$*, then* $x_T = x^*$ *and*

$$J^*(x) = \begin{cases} J_V^m(x) & x < x^* \\ \underline{f}_{x^*}(x) & x \geq x^*, \end{cases}$$

where $m = \int_V f(x)dx + \Delta$.

Panel (b) in Figure 6 graphically illustrates the strategic justification J^* for the part (b) of the above proposition. As before, the expert inflates the strength of the evidence weaker than x^*, and deflates the strength of the evidence stronger than x^*. And, such a misrepresentation harms a client c with borderline signal realizations $x_c \in [f^{-1}(c-b), f^{-1}(c))$ if (c, x_c) does not belong to triangular regions located at kinks of J^*. These triangular regions highlight client and signal realization pairs (c, x_c) that benefit from the the auditor's anecdotes.

Expert's strategic justification suggests that providing the auditor with more anecdotes on signal realizations larger than x^* cannot harm clients. Such anecdotes change the strategic justification on set V^c without changing the expert's actions. Moreover, they weakly reduce the mass $\Delta = \{D(K,f), S(K,f)\}$ by weakly reducing the supply $S(K,f)$. Thus, strategic justification with additional anecdotes comes closer to the actual density on signal realizations smaller than x^*.

The analysis is more nuanced when all the additional anecdotes are smaller than x^*. These anecdotes reduce the demand $D(K,f)$ which may or may not reduce the effective demand $\Delta = \{D(K,f), S(K,f)\}$. However, the kinks at these additional anecdotes change how the new mass m'

is spread across signal realizations in set V. Thus, the net effect depends on a reduction in Δ and the shape of new strategic distribution $J_V^{m'}$. The next subsection analyzes this setting.

5.2 Harmful expertise

The set K' contains more anecdotes than set K and all the additional anecdotes are smaller than x^*. Thus $K_{x^*} = K'_{x^*}$, where $K_{x^*} = \{x_j \in K : x_j \geq x^*\}$. The latter assumption implies that $S(K, f) = S(K', f)$. The comparison of aggregate payoffs of clients for auditors K and K' depends on whether the expert is demand or supply-constrained.

Proposition 3. *(a) If $D(K, f) \leq S(K, f)$, then aggregate payoff of clients for the auditor K' is weakly larger than that for auditor K. (b) If $D(K', f) > S(K', f)$, then aggregate payoff of clients for the auditor K is weakly larger than that for auditor K'.*

Proof. For part (a), the expert is demand-constrained with auditor K, i.e., $D(K, f) \leq S(K, f)$. Observe that $D(K', f) \leq D(K, f)$, the new anecdotes on $[0, x^*)$ locally constrain the expert and weakly reduce the demand of misrepresentation. Thus, the expert is demand-constrained, $D(K', f) < S(K', f)$, even for auditor K'. Observing that the strategic justification for the auditor K' pointwise dominates that for the auditor K completes the proof.

For part (b), observe that an expert who is supply-constrained for the auditor K' is also supply-constrained for auditor K. Thus, in both cases, the expert's strategic justification integrates to the same number $\int_0^{x^*} f(x)dx + S(K, f)$ on the interval $[0, x^*)$. However, both these strategic justifications distribute this mass differently on $[0, x^*)$ (due additional anecdotes in K'). To compare the clients' aggregate payoff Π_{Agg} across these two strategic justifications, we rewrite the expression as

$$\begin{aligned}\Pi_{Agg}(K, f) &= -\frac{1}{2\bar{c}} \int_0^{x^*} \left[J^*(x) \left(\frac{J^*(x)}{2} \right) + (\bar{c} - J^*(x)) f(x) \right] dx - \frac{\bar{c}(1 - x^*)}{4} \\ &=_{(a)} \Pi_E(K, f) - \frac{b}{2\bar{c}} \int_0^{x^*} J^*(x) dx - \frac{b(1 - x^*)}{2}.\end{aligned}$$

Equality (a) follows from adding and subtracting the last two terms. Since the expert earns weakly larger payoff when the auditor has less anecdotes ($\Pi_E(K, f) \geq \Pi_E(K', f)$), we obtain the ranking of clients' aggregate payoffs proposed in the statement of the proposition. □

To see the intuition for the part (b), consider the direct effect of the additional anecdotes. They restrict the expert's ability to inflate the justification above the actual density f but only locally. The indirect effect comes through a change in the expert's strategic justification. The expert's new strategic justification moves away from actual density f at signal realizations farther from the anecdotes. This indirect effect dominates the direct effect and reduces the clients' aggregate payoff.

The above analysis considered all cases except $D(K', f) < S(K, f) < D(K, f)$. In this case, the expert is supply-constrained with the auditor K but becomes demand-constrained for the auditor K'. In this case, the mass m spread on signal realizations smaller than x^* and its distribution J_V^m changes. Since both factors have opposite consequences, the comparison depends on the exact details of K, K', and f.

6 Strategic consistency

In previous sections, the expert passed the auditor's scrutiny as long as the expert's justification J was a non-decreasing, bounded, and atomless density function such that $J(x) = f(x)$ for $x \in K$. This section considers an auditor who reasons strategically and uses a stringent criterion for evaluating the feasibility of justifications. Such an auditor requires the expert's justifications J to satisfy the following additional property.

Definition 5 (Strategic Consistency)**.** *If J is the actual density, then an expert with payoff parameter b must report a justification that coincides with J almost everywhere on* $[0, 1]$.

Strategic consistency avoids suspicion of the auditor who is aware of the qualitative properties of the strategic justification and knows that a self-interested expert has incentives to deviate from the actual density f. Proposition 1S (where S stands for strategic consistency) below shows that Proposition 1 extends with this stringent definition of feasible justification if the auditor knows that the expert's knowledge is always valuable on $[0, 1]$.

Proposition 1S. *Suppose that the expert's knowledge is always valuable on* $[0, 1]$*, and the auditor is aware of it. Then the expert's strategic justification coincides with f almost everywhere on* $[0, 1]$.

Proof. If the expert's knowledge is always valuable on $[0, 1]$ then by Proposition 1 the expert's strategic justification equals f almost everywhere. Since the auditor knows that the expert's knowledge is always valuable on $[0, 1]$, he also knows that the reported justification coincides with f almost everywhere. Thus, the strategic justification satisfies strategic consistency. □

We make the following assumptions, in addition to Assumption 1A, to analyze the case when the expert's knowledge is not always valuable on $[0, 1]$.

Assumption 2. *f is continuous at x^* and $x^* \in K$.*

The first part of the above assumption implies that $f(x^*) = \bar{c}$. Moreover, $V = \{x : f(x) < \bar{c}\} = [0, x^*)$. The second part implies that the auditor knows $f(x^*) = \bar{c}$. Such knowledge of the auditor could result from knowing the client-optimal decision rule for the highest client $\bar{c}$. The following assumption, says that for any given density f and payoff parameter $b > 0$, the auditor's anecdotes in K are close enough so that $\min\{f(x) + b, \overline{f}_{x^*}(x)\} = \overline{f}_{x^*}(x)$ for $x < x^*$.

Assumption 3. *For any signal realization $x < x^*$, $\overline{f}_{x^*}(x) - \underline{f}_{x^*}(x) \leq b$.*

Proposition 2S below shows that the qualitative properties of the expert's strategic justification continue to hold even with the stringent criterion of strategic consistency.

Proposition 2S. *Suppose the expert's knowledge is not always valuable on $[0, 1]$ and f satisfies Assumptions 1A, 2 and 3. Then the expert's strategic justification J^* lies above the actual density f for signal realizations smaller than x^* and otherwise lies weakly below f. In particular,*
(a) If $D(K, f) \leq S(K, f)$, then $J^*(x) = \overline{f}_{x^*}(x)$ for $x < x^*$.
(b) If $D(K, f) > S(K, f)$, then

$$J^*(x) = \begin{cases} J_V^m(x) & x < x^* \\ \underline{f}_{x^*}(x) & x \geq x^*, \end{cases}$$

where $m = \int_V f(x)dx + \Delta$.

Proof. Suppose, $D(K, f) \leq S(K, f)$. By Proposition 2A, the expert's strategic justification is such that $J^*(x) = \min\{f(x) + b, \overline{f}_{x^*}(x)\}$ for $x < \hat{x}$, where $\hat{x}$ is the smallest signal realizations satisfying $\min\{f(x) + b, \overline{f}_{x^*}(x)\} = \bar{c}$. Assumption 3 implies that $\min\{f(x) + b, \overline{f}_{x^*}(x)\} = \overline{f}_{x^*}(x)$. Thus, $J^*(x) = \overline{f}_{x^*}(x)$ for $x < \hat{x}$. For $x \in [\hat{x}, x^*)$, we must have $J^*(x) = \overline{f}_{x^*}(x) = \bar{c}$. To see why, observe that the auditor knows that $f(x^*) = \bar{c}$ (Assumption 2). Therefore, the expert cannot report a justification that is strictly above $\bar{c}$ for $x < x^*$ without violating the monotonicity property. To complete the proof it suffices to show that any justification with the proposed form satisfies strategic consistency. Suppose J^* is the actual density. Then,

$$D(K, J^*) = \int_0^{x^*} \left(\min\{J^*(x) + b, \overline{f}_{x^*}(x)\} - J^*(x)\right) dx = \int_0^{x^*} \left(\overline{f}_{x^*}(x) - \overline{f}_{x^*}(x)\right) dx = 0.$$

Thus, the expert cannot inflate the justification above J^* without violating feasibility. Therefore, have no incentive to deviate from J^*.

Consider the case $D(K, f) > S(K, f)$. Then by Proposition 2A, the expert's strategic justification is such that $\underline{f}_{x^*}(x)$ for $x \geq x^*$. We complete the proof by by showing that J^* with the proposed form satisfies strategic consistency. Since the auditor knows that $f(x^*) = \bar{c}$ (Assumption 2), the expert cannot report a justification that is strictly below $\underline{f}_{x^*}(x)$ for $x \geq x^*$. Thus, $S(K, J^*) = 0$ and the expert has no incentive to deviate from J^*. □

7 Conclusion

This paper considered a self-interested expert who was required to justify her decisions to an auditor. Even though the auditor had very little expertise, he could verify the logical consistency of the expert's justifications. It showed that providing logically coherent justification for past decisions sometimes is sufficient to force the expert to act in the clients' best interest. Otherwise, the expert devised a justification that manipulated the strength of the evidence and made the expert's selfish actions appear client-optimal. For the latter case, the paper established that increasing the auditor's expertise, which only locally improved the auditor's interpretation skills, could backfire and reduce the aggregate payoff of clients.

This paper proposed a tractable framework to study incentive conflict in a setting of asymmetric expertise. It exhibited how thinking about expertise as a non-verifiable function, which satisfies specific logical consistency properties, can give new and exciting results. Moreover, it shed light on the workings of the institutional arrangement of fiduciary duty. We believe tools and modeling ideas developed in this paper could prove beneficial in understanding the role of justifications and stories in organizations. We plan to explore these ideas in a separate project.

Appendix A states definitions of functions $\overline{f}$ and $\underline{f}$. Appendix B lists mathematical concepts that are essential for the analysis and proofs. An extensive treatment of these concepts can be found in standard text-books like Royden and Fitzpatrick (1988) and Ok (2011). Appendix C contains proofs that do not appear in the main text. Appendix D contains technical lemmas used throughout the Appendix C.

A Definitions of bounds

The lower bound $\underline{f}$ and the upper bound $\overline{f}$ are defined below for $x \in [0, 1]$.

$$\underline{f}(x) = \begin{cases} 0 & 0 \leq x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ . & \\ f(x_{k-1}) & x_{k-1} \leq x < x_k \\ f(x_k) & x = x_k \end{cases} \qquad \overline{f}(x) = \begin{cases} f(x_1) & 0 \leq x \leq x_1 \\ f(x_2) & x_1 < x \leq x_2 \\ . & \\ f(x_{k-1}) & x_{k-2} < x \leq x_{k-1} \\ f(x_k) & x_{k-1} < x \leq x_k \end{cases} .$$

B Mathematical preliminaries

Let $1 \leq p < \infty$. A Lebesgue measurable function h defined on $[\underline{x}, \overline{x}]$ is said to belong to the space $L^p = L^p[\underline{x}, \overline{x}]$ if $\int_{\underline{x}}^{\overline{x}} |h(x)|^p dx < \infty$. L^p space is a complete normed linear space, i.e., a Banach space with the norm

$$\|h\|_p = \left\{ \int_{\underline{x}}^{\overline{x}} |h(x)|^p dx \right\}^{1/p}.$$

Similarly, L^∞ is the space of bounded Lebesgue measurable (henceforth 'measurable') functions on $[\underline{x}, \overline{x}]$. The associated norm is

$$\|h\|_\infty = \inf\{M : \mu\{x : h(x) > M\} = 0\},$$

where μ is the Lebesgue measure. Unless mentioned otherwise, we consider convergence and continuity with respect to the topology of the norm $\|\cdot\|_p$.

Fréchet derivatives: Consider a function $\phi : L^p \to V$, where V is a Banach space with norm $\|\cdot\|_q$ and $1 \leq q < \infty$. Function ϕ is Fréchet differentiable at $J \in L^p$ if there exists a bounded linear operator ${\phi'}_J : L^p \to V$ such that

$$\lim_{t \to 0} \frac{\|\phi(J + t) - \phi(J) - {\phi'}_J(t)\|_q}{\|t\|_p} = 0,$$

where ϕ'_J is bounded if there is constant M such that $\|\phi'_J(f)\|_q \leq M\|f\|_p$. The linear operator ϕ'_J is called the Fréchet derivative of ϕ at J. Moreover, Fréchet differentiability of ϕ at J implies continuity of ϕ at J. We say that function ϕ is Fréchet differentiable on L^p if it is differentiable at every $J \in L^p$. In this case, ϕ' denotes the Fréchet derivative. We will use the following result to establish the continuity of ϕ'.

Result 1. *If ϕ is Fréchet differentiable at every $J \in L^p$, then ϕ is continuously Fréchet differentiable.*

The above result follows from well-known theorem of Banach: A linear operator on a normed linear space is continuous iff it is bounded.

Dual representation: Let $X \subset L^p$, where $1 \leq p < \infty$. The space of all bounded linear functional F on X is termed as a dual spaces of X, and is denoted as X^*. Define

Result 2. *Let $X = L^p$ and q be such that $1/p + 1/q = 1$. Then each function $g \in L^q$ defines a bounded linear functional F on L^p*

$$F(\phi) = \int_{\underline{x}}^{\overline{x}} \phi(x)g(x)dx.$$

Moreover, $\|F\| = \|g\|_q$.

C Proofs not in the main text

Proof of Theorem 1. Fix a justifiable decision rule d. The proof constructs a feasible justification J such that $a_c(d, x_c) = a_c(d^*_J, x_c)$ for every (c, x_c) pair. Consider the function $\hat{J}(x) = J_{D(x)}(x)$, where $J_{D(x)}(x)$ was defined in the discussion leading to Observation 2. Let $V = \{x : \hat{J}(x) < \bar{c}\}$.

First, we state some properties of function $\hat{J}$ in relation to the decision rule d. Observation 2 implies that the decision rule d satisfies

$$a_c(d, x) = \begin{cases} H & x \in V \text{ and } c \leq \hat{J}(x) \\ L & x \in V \text{ and } c > \hat{J}(x) \\ H & x \in V^c \end{cases}.$$

Lemma 1 exhibits that $\hat{J}(x)$ is non-decreasing on V, where $V = [0, x^*)$ or $V = [0, x^*]$ and $x^* \leq 1$. Since $\hat{J}(x)$ is constructed from the point-by-point piecing of bounded and atomless functions, it also satisfies these properties.

Suppose $x^* = 1$. By properties stated in the previous paragraph, the function $J(x) = \hat{J}(x)$ is a non-decreasing, bounded, and atomless function that satisfies $a_c(d, x_c) = a_c(d^*_J, x_c)$ for every (c, x_c)

pair. Thus, to complete the proof it suffices to show that $\int_0^1 \hat{J}(x)dx = 1$. Lemma 2 establishes the same.

Suppose $x^* < 1$. We define the feasible density J as

$$J(x) = \begin{cases} \hat{J}(x) & x \in V \\ g(x) & x \in V^c \end{cases}.$$

Here g is a feasible density such that $g(x) \geq \overline{c}$ for $x \in V^c$, and $\int_{V^c} g(x)dx = 1 - \int_V \hat{J}(x)dx$. Such a function g exists as per Lemma 3. Observe that $J(x)$ is a feasible density function as it is a non-decreasing, bounded, and atomless function that integrates to 1. The proof is completed by observing that for any given pair (c, x_c), we have $a_c(d, x_c) = a_c(d^*_J, x_c)$. □

Proof of Observation 3. Suppose that the expert follows a justifiable decision rule d. Then, the expert's total expected payoff is

$$\begin{aligned}
\Pi_E(d, f) &= \frac{1}{\overline{c}} \int_0^{\overline{c}} \pi_E(c, f, d)dc \\
&=_{(a)} -\frac{1}{2\overline{c}} \int_0^{\overline{c}} \int_0^1 \left[(c-b)d(c,x) + f(x)(1-d(c,x))\right] dx\ dc \\
&=_{(b)} -\frac{1}{2\overline{c}} \int_0^1 \left[\int_0^{\overline{c}} (c-b)d(c,x)dc + \int_0^{\overline{c}} f(x)(1-d(c,x))dc\right] dx \\
&=_{(c)} -\frac{1}{2\overline{c}} \int_0^1 \left[\int_0^{min\{J(x),\overline{c}\}} (c-b)dc + \int_{min\{J(x),\overline{c}\}}^{\overline{c}} f(x)dc\right] dx \\
&= \quad -\frac{1}{2\overline{c}} \int_0^1 \left[J_{\overline{c}}(x)\left(\frac{J_{\overline{c}}(x)}{2} - b\right) + (\overline{c} - J_{\overline{c}}(x))f(x)\right] dx.
\end{aligned}$$

Equality (a) is obtained by substituting equation (1) from Section 2. Whereas, equality (b) is obtained by changing the order of integral. Such a change of order is a valid operation because signal realization x_c is independent of the identifier c as well as signal realizations of other clients. By Theorem 1, there exists a feasible density J such that the expert takes action H iff $c \leq J(x)$. Thus, Equality (c) follows. Finally, by writing $J_{\overline{c}}(x) = min\{J(x), \overline{c}\}$ and omitting the leading constant, the objective function in the statement of the observation is obtained. □

Proof of Theorem 2. Following the discussion leading to the statement of Theorem 2, the expert's problem can be written as

$$\max_{J} \; -\int_{\underline{x}}^{\overline{x}} \left[J_{\overline{c}}(x)\left(\frac{J_{\overline{c}}(x)}{2} - b\right) + (\overline{c} - J_{\overline{c}}(x))f(x)\right] dx$$

$$\text{subject to} \quad J \;\text{ is a non-decreasing, bounded, and atomless function,} \qquad \text{(monotonicity)}$$

$$\int_{\underline{x}}^{\overline{x}} J(x)dx = m, \qquad \text{(aggregate consistency)}$$

where $J_{\overline{c}}(x) = min\{J(x), \overline{c}\}$ and $\int_{\underline{x}}^{\overline{x}} f(x)dx \leq m \leq \int_{\underline{x}}^{\overline{x}} \min\{f(x) + b, \overline{c}\}dx$. Suppose m equals the upper bound. In this case, the expert has the necessary mass to inflate the justification at all signal realizations and follow the unconstrained optimal decision rule for all feasible (c, x) pairs. Observing that the proposed J^* in the theorem statement coincides with the unconstrained optimal decision rule for $\lambda = 0$ completes the proof.

Suppose, $\int_{\underline{x}}^{\overline{x}} f(x)dx \leq m < \int_{\underline{x}}^{\overline{x}} \min\{f(x) + b, \overline{c}\}dx$. Before solving the expert's problem, we make the following changes to it. Consider a feasible justification J such that $J(x) > \overline{c}$ on a non-empty interval $I \subset [\underline{x}, \overline{x}]$ of non-zero length. Since m is strictly smaller than the upper bound, there exist a positive mass of signal realizations where $J(x) < \min\{f(x) + b, \overline{c}\}$. Reducing $J(x)$ to $\overline{c}$ on I and using the mass $\int_I J(x)dx - \overline{c}|I|$ to inflate the justification at other signal realizations allows the expert to move closer to her unconstrained optimal decision rule and increases her payoff. Therefore, J is dominated. This observation implies that imposing the constraint $J(x) \leq \overline{c}$ in the above problem does not change the optimal solution. Moreover, the new constraint allows us to write $J_{\overline{c}}(x) = J(x)$ in the objective function. Finally, relaxing the constraint "J is non-decreasing and atomless function" and incorporating the constraint "J is a bounded function" through the set $B = \{J : |J(x)| \leq M \;\; \forall x \in [0, 1]\}$, we obtain

$$\min_{J \in L^1 \cap B} H(J) := \int_{\underline{x}}^{\overline{x}} \underbrace{\left[J(x)\left(\frac{J(x)}{2} - b\right) + (\overline{c} - J(x))f(x)\right]}_{h(J)} dx \qquad \text{(MP)}$$

$$\text{subject to} \quad u(J) := J(x) - \overline{c} \leq 0 \quad \forall\, x \in [\underline{x}, \overline{x}], \qquad \text{(affine constraint)}$$

$$e(J) := m - \int_{\underline{x}}^{\overline{x}} J(x)dx = 0. \qquad \text{(aggregate consistency)}$$

Notice that we have stated the above problem as a minimization problem. In the remaining proof, we use theory of Lagrange multipliers on functional spaces to solve the relaxed problem (MP) and show that the solution is a "non-decreasing and atomless function."

Let $\mathcal{Y}$ denote the set of viable functions that satisfy the constraints of problem (MP). The solution proposed in the statement of the Theorem 2 satisfies $J^* \in \mathcal{Y}$. Observe that the constraints of set $\mathcal{Y}$ are linear, and the set $\mathcal{Y}$ is convex. Moreover, $H(J)$ is continuous (Lemma 7) and strictly convex (Lemma 6) in J. Thus, for any $J \in \mathcal{Y}$ to be the global solution of the problem (MP), the first-order Karush-Kuhn-Tucker (KKT) conditions are both necessary and sufficient. The remaining proof shows that J^* satisfies KKT conditions and that the problem (MP) has only one global minimizer in the almost everywhere sense. Thus, J^* is an unique solution.

Observe that $H : L^1 \cap B \to \mathbb{R}$ is a functional. Let $K = \{v \in L^2 : v(x) \leq 0\}$ be a closed convex cone. Then the function $u : L^1 \cap B \to K$ represents an infinite-dimensional affine constraint. For any given $J \in \mathcal{Y}$, we denote the value of constraint u at a given x by $u(J)(x)$. Finally, the equality constraint can be thought as $e : L^1 \cap B \to \{0\}$. Let $K^+ = \{v \in L^2 : v(x) \geq 0\}$ and define a Lagrangian function $\mathcal{L} : L^1 \times \mathbb{R} \times K^+ \to \mathbb{R}$ such that

$$\mathcal{L}(J, \lambda, \eta_u) = H(J) + \lambda(e(J)) + \int_{\underline{x}}^{\overline{x}} \eta_u(x)u(J)(x)dx.$$

Lemma 7 states that functionals H, e, and $U = \int_{\underline{x}}^{\overline{x}} \eta_u(x)u(J)(x)dx$ are continuously Fréchet differentiable on $L^1 \cap B$, a closed and convex set. Their Fréchet derivatives are $H'_J(\phi) = \int_{\underline{x}}^{\overline{x}}(J(x) - b - f(x))\phi(x)dx$, $e'_J(\phi) = -\int_{\underline{x}}^{\overline{x}} \phi(x)dx$, and $U'_J(\phi) = \int_{\underline{x}}^{\overline{x}} \eta_u(x)\phi(x)dx$. Since J^* satisfies certain regularity conditions (Lemma 8), Theorem 1.56 in Hinze et al. (2008) implies that there exist Lagrange multipliers $\lambda \in \mathbb{R}$ and $\eta_u \in K^+$ such that

$$\int_{\underline{x}}^{\overline{x}} \underbrace{[J^*(x) - b - f(x) - \lambda + \eta_u(x)]}_{g(x)} \phi(x)dx = 0 \qquad \text{for any } \phi \in L^1 \tag{FOC}$$

$$e(J^*) = 0 \text{ and } \ u(J^*)(x) \leq 0 \ \ \forall\ x \in [\underline{x}, \overline{x}] \tag{Original Constraints}$$

$$\eta_u(x)\,(u(J^*)(x)) = 0 \text{ a.e. } x \in [\underline{x}, \overline{x}]. \tag{Complementary Slackness}$$

(FOC) gives the first-order stationarity condition obtained using Fréchet derivatives of the functional $\mathcal{L}$. It can be restated as

$$J^*(x) - b - f(x) - \lambda + \eta_u(x) = 0 \text{ a.e. } x \in [\underline{x}, \overline{x}]. \tag{FOC-M}$$

To show that J^* is indeed a solution, it remains to be shown that it satisfies the above necessary conditions. First, notice that J^* satisfies (Original Constraints). Let $\eta_u(x) = \max\{f(x) + b + \lambda - J^*(x), 0\}$. Observe that $\eta_u(x) \geq 0$ for any $x \in [\underline{x}, \overline{x}]$ and thus belongs to K^+. Moreover, J^* and η_u satisfy (Complementary Slackness) and the stationarity condition (FOC-M). Observe that the constraints on m imply that $-b \leq \lambda \leq 0$. Finally, the strict convexity of $H(J)$ implies that J^* is a unique global minimizer (Lemma 9).

We complete the proof by observing that J^* satisfies the relaxed constraint as it is a non-decreasing function. Moreover, it is obtained by piecing $f(x)+b+\lambda$ and $\bar{c}$ that do not have jumps of positive Lebesgue measure. $\square$

Proof of Theorem 2A. Using Observation 4 and the discussion leading to the statement of Theorem 2A, the expert's problem can be written as

$$\max_J \; -\int_{\underline{x}}^{\overline{x}} \left[J_{\bar{c}}(x)\left(\frac{J_{\bar{c}}(x)}{2} - b\right) + (\bar{c} - J_{\bar{c}}(x))f(x)\right] dx$$

$$\begin{aligned} \text{subject to } \; & J \text{ is a non-decreasing, bounded, and atomless function,} && \text{(monotonicity)} \\ & \underline{f}(x) \leq J(x) \leq \overline{f}(x) \;\; \text{for } x \in [\underline{x}, \overline{x}], && \text{(local consistency)} \\ & \int_{\underline{x}}^{\overline{x}} J(x)dx = m, && \text{(aggregate consistency)} \end{aligned}$$

where $J_{\bar{c}}(x) = min\{J(x), \bar{c}\}$.

First, we argue that replacing the constraint $J(x) \leq \overline{f}(x)$ with the constraint $J(x) \leq \min\{\overline{f}(x), \bar{c}\}$ does not change the solution of the above problem. Suppose $f(\overline{x}) \leq \bar{c}$. Then, $\overline{f}(x) \leq \bar{c}$ for all $x \in [\underline{x}, \overline{x}]$ and the required result follows. If $f(\overline{x}) > \bar{c}$, then consider a feasible justification J such that $J(x) > \bar{c}$ on a non-empty interval $I \subset [\underline{x}, \overline{x}]$ of non-zero length. Since $m \leq \int_{\underline{x}}^{\overline{x}} \min\{f(x) + b, \overline{f}(x), \bar{c}\}dx$, there exist a positive measure of signal realizations where $J(x) < \min\{f(x)+b, \overline{f}(x), \bar{c}\}$. Reducing $J(x)$ to $\min\{\underline{f}(x), \bar{c}\}$ on I and using the mass $\int_I \left(J(x) - \min\{\underline{f}(x), \bar{c}\}\right) dx$ to inflate the justification at aforementioned signal realizations allows the expert to move closer to her unconstrained optimal decision rule and increases her payoff. Therefore, J is dominated. In other words, replacing the constraint $J(x) \leq \overline{f}(x)$ with the constraint $J(x) \leq \min\{\overline{f}(x), \bar{c}\}$ does not change the solution of the above problem. Moreover, it also allows us to write $J_{\bar{c}}(x) = J(x)$ in the objective function.

Using the discussion in the previous paragraph, we write the expert's problem as

$$\max_J \; -\int_{\underline{x}}^{\overline{x}} \left[J(x)\left(\frac{J(x)}{2} - b\right) + (\bar{c} - J(x))f(x)\right] dx$$

$$\begin{aligned} \text{subject to } \; & J \text{ is a non-decreasing, bounded, and atomless function,} && \text{(monotonicity)} \\ & \underline{f}(x) \leq J(x) \leq \min\{\overline{f}(x), \bar{c}\} \;\; \text{for } x \in [\underline{x}, \overline{x}], && \text{(modified local consistency)} \\ & \int_{\underline{x}}^{\overline{x}} J(x)dx = m. && \text{(aggregate consistency)} \end{aligned}$$

To solve this problem, we relax the constraint $J(x) \geq \underline{f}(x)$ for all $x \in [\underline{x}, \overline{x}]$. The resulting problem is analogous to the problem (MP) with the difference that the affine constraint is now changed to

$J(x) \le \min\{\overline{f}(x), \overline{c}\}$. Thus, the proof of Theorem 2 can be used to show that J^* proposed in the statement of the Theorem 2A is a solution. To complete the proof, observe that this J^* respects the relaxed constraint $J(x) \ge \underline{f}(x)$. $\square$

Proof of Proposition 2A. Part (a) holds by a reasoning similar to the proof of part (a) in Proposition 2. The only difference is that the strategic justification in the current case also respects the local consistency constraint.

Now consider the case, $D(K, f) > S(K, f)$. Let J be a feasible justification such that $J(x) > \underline{f}_{x^*}(x)$ on a non-empty interval $I \subset V^c$ of non-zero length. Since the expert is supply-constrained, there exist a positive mass of signal realizations in V where $J(x) < \min\{f(x)+b, \overline{f}_{x^*}(x)\}$. Reducing $J(x)$ to $\underline{f}_{x^*}(x)$ on I and using the mass $\int_I \Big(J(x) - \underline{f}_{x^*}(x)\Big)\, dx > 0$ to inflate the justification on aforementioned signal realizations in V increases the expert's payoff. Therefore, J is dominated. This observation implies that the strategic justification must satisfy $J^*(x) \le \underline{f}_{x^*}(x)$ for all $x \in V^c$.

Suppose, $x^* \in (x_{j-1}, x_j]$ where $j \le k$ and $x_{j-1}, x_j \in K$. If $j = 1$ then $x_{j-1} = 0$. The auditor knows that any feasible density cannot be smaller than $\underline{f}_{x^*} = \underline{f}$ for $x \in [x_j, 1]$. Thus, $J^*(x) = \underline{f}_{x^*}(x)$ for $x \in [x_j, 1]$. Now consider $x \in [x^*, x_j)$. For these signal realizations, $\underline{f}_{x^*}(x) = \overline{c}$. Thus, we can conclude that strategic justification must satisfy $J^*(x) \le \overline{c}$ for $x \in [0, x_j)$. Let $m = 1 - \int_{x_j}^1 \underline{f}_{x^*}(x)dx$. Observe that

$$\int_0^{x_j} f(x)dx \le m \le \int_0^{x_j} \min\{f(x) + b, \overline{f}(x), \overline{c}\}dx.$$

The left inequality holds because $\int_{x_j}^1 \underline{f}_{x^*}(x)dx \le \int_{x_j}^1 f(x)dx$. The right inequality is true because the expert is supply-constrained.

Lemma 10 implies that the strategic justification has the form $J^*(x) = \min\{f(x)+b+\lambda, \overline{f}(x), \overline{c}\}$ on the interval $x \in [0, x_j)$, where $-b \le \lambda \le 0$ and $\int_0^{x_j} J^*(x)dx = 1 - \int_{x_j}^1 \underline{f}_{x^*}(x)dx$. By arguments similar to those in the last two paragraphs of the proof of Proposition 2, we obtain the desired result. $\square$

D Proofs of lemmas

Lemma 1. *$\hat{J}(x)$ is non-decreasing on V. Moreover, set V is either a right-closed $[0, x^*]$ or a right-open $[0, x^*)$ interval, where $x^* \le 1$.*

Proof. To prove the first part, assume to the contrary that $\hat{J}(x)$ is a not non-decreasing function on V. Then there exist $x_1, x_2 \in V$ such that $x_1 < x_2$ and $\hat{J}(x_1) > \hat{J}(x_2)$. We use the existence of such x_1, x_2 to show that the decision rule d is not justifiable, thus yielding a contradiction.

Let $c = \hat{J}(x_1)$ and $c' = \hat{J}(x_2) + \epsilon$, where $\epsilon < \hat{J}(x_1) - \hat{J}(x_2)$. Consider any data D containing points $(x_c = x_1, a_c(d, x_1))$ and $(x_{c'} = x_2, a_{c'}(d, x_2))$. By definition of $\hat{J}$, the decision rule d is such that $a_c(d, x_1) = H$ whereas $a_{c'}(d, x_2) = L$. Thus, any feasible justification J for data D must satisfy $J(x_1) \geq \hat{J}(x_1)$ and $J(x_2) < \hat{J}(x_2) + \epsilon$. These two inequalities imply that J is not a non-decreasing function, which contradicts the fact that the decision rule d is justifiable.

To prove the second part, again assume for a contradiction. Thus, there exist $x_1 < x_2$ such that $x_1 \in V^c$, $x_2 \in V$ and $\hat{J}(x_1) > \hat{J}(x_2)$. Using argument similar to the first part, we show that decision rule d is not justifiable. □

Lemma 2. *If* $x^* = 1$ *then* $\int_0^1 \hat{J}(x)dx = 1$.

Proof. Assume for a contradiction, i.e., $\int_0^1 \hat{J}(x)dx \neq 1$. First, consider the case $\int_0^1 \hat{J}(x)dx > 1$. We construct data $D = \{x_c, a_c(d, x_c)\}_{c \in \mathcal{C}}$ that originates from the decision rule d and is not justifiable. We start by considering a function $\tilde{J}(x)$, the inverse of which will be used to construct the desired data D.

Let $V_+ = \{x \in V : \hat{J}(x) > 0\}$ be the set of signal realizations where $\hat{J}(x)$ is strictly positive. The set $\overline{V_+} = [s, 1]$ denotes the closure of set V_+, where $s \geq 0$. Let ϵ be such that $0 < \epsilon < \int_0^1 \hat{J}(x)dx - 1$. By Lemma 4, there exists a function $\tilde{J}(x)$ defined on the set $(s, 1]$ such that

1. $\tilde{J}(x)$ is strictly increasing and continuous on $(s, 1]$ with $\tilde{J}(1) = \tilde{J}(1^-)$,
2. $\tilde{J}(x) \leq \hat{J}(x)$ for all $x \in (s, 1]$,
3. $\int_s^1 \tilde{J}(x)dx > \int_s^1 \hat{J}(x)dx - \epsilon > 1$,

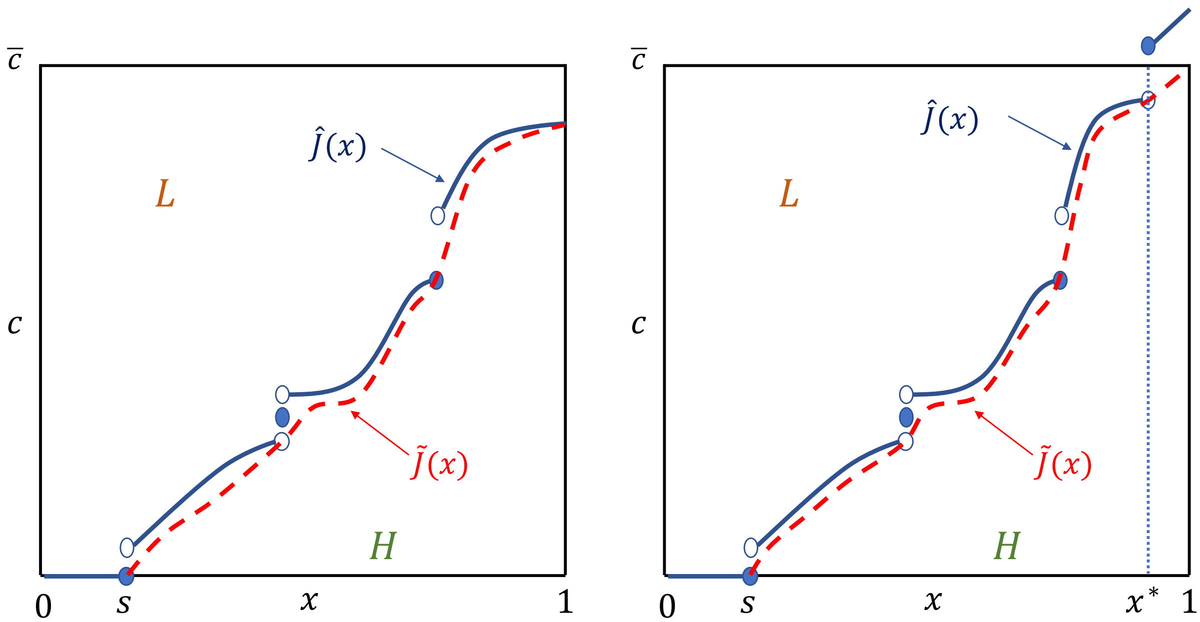
where $\tilde{J}(x^-)$ denotes the left-hand limit of $\tilde{J}(\cdot)$ at x. Panel (a) in Figure 7, qualitatively illustrates the function $\tilde{J}(x)$ on the interval $(s, 1]$. The definition of $\tilde{J}$ can be extended to all $x \in [0, 1]$ by letting $\tilde{J}(x) = \hat{J}(x)$ for all $x \in [0, s)$ and $\tilde{J}(s) = \min\{\hat{J}(s), \tilde{J}(s^+)\}$. Notice that the extended $\tilde{J}$ satisfies Property 2 and Property 3 on $[0, 1]$.

Using the function $\tilde{J}$, we construct data D that is not justifiable. Consider

$$D = \begin{cases} (s, a_c(d, s))) & c \in [0, \tilde{J}(s^+)] \\ \left(\tilde{J}^{-1}(c), a_c(d, \tilde{J}^{-1}(c))\right) & c \in R = (\tilde{J}(s^+), \tilde{J}(1)] \\ (1, a_c(d, 1)) & c \in (\tilde{J}(1), \overline{c}] \end{cases} .$$

Property 1 implies that $\tilde{J}(x)$ is strictly increasing and continuous on $(s, 1]$, and therefore the inverse $\tilde{J}^{-1}(c)$ is well-defined for all $c \in R$.

To see why data D is not justifiable, fix a client $c \in R = (\tilde{J}(s^+), \tilde{J}(1)]$. For the client c, the signal realization x_c is such that $c = \tilde{J}(x_c) \leq \hat{J}(x_c)$. The last inequality follows from Property



Panel (a): $\tilde{J}(x)$ in Lemma 2 　　 Panel (b): $\tilde{J}(x)$ in Lemma 3

Figure 7: Graphical illustration of function $\tilde{J}$ in Lemmas 2 and 3.

2. Therefore, the action associated with client c in data D is $a_c(d, x_c) = H$. In other words, any feasible density function J that justifies D must be weakly larger than $\tilde{J}(x)$. Thus, $\int_0^1 J(x)dx \geq \int_0^1 \tilde{J}(x)dx > \int_0^1 \hat{J}(x)dx - \epsilon > 1$, where the second inequality follows from Property 3. This inequality contradicts that the justification J is feasible and that the decision rule d is justifiable. Proof for the case $\int_0^1 \hat{J}(x)dx < 1$ mirrors the above argument. □

Lemma 3. *If $x^* < 1$ then there exists a feasible density g such that $g(x) \geq \overline{c}$ for $x \in V^c$, and $\int_0^1 J(x)dx = 1$, where*

$$J(x) = \begin{cases} \hat{J}(x) & x \in V \\ g(x) & x \in V^c \end{cases}.$$

Proof. Assume for contradiction. Then for all feasible densities g such that $g(x) \geq \overline{c}$ for $x \in V^c$ we must have $\int_0^1 J(x)dx > 1$. In particular, $\int_{V^c} \overline{c}dx > 1 - \int_V \hat{J}(x)dx$. We use this latter inequality to construct a data realization D that is not justifiable.

Let $V_+ = \{x \in V : \hat{J}(x) > 0\}$ be the set of signal realizations where $\hat{J}(x)$ is strictly positive. The set $\overline{V_+} = [s, x^*]$ denotes the closure of set V_+, where $s \geq 0$. Let $\epsilon > 0$ be such that $\epsilon < \int_0^{x^*} \hat{J}(x)dx + \int_{x^*}^1 \overline{c}dx - 1$. By Lemma 4, there exists a function $\tilde{J}(x)$ defined on the set $(s, x^*]$ such that

1. $\tilde{J}(x)$ is strictly increasing and continuous on $(s, x^*]$ with $\tilde{J}(x^*) = \tilde{J}((x^*)^-)$,
2. $\tilde{J}(x) \leq \hat{J}(x)$ for all $x \in (s, x^*]$,
3. $\int_s^{x^*} \tilde{J}(x)dx > \int_s^{x^*} \hat{J}(x)dx - \epsilon/2$.

The definition of $\tilde{J}$ can be extended to all $x \in [0, x^*]$ by letting $\tilde{J}(x) = \hat{J}(x)$ for all $x \in [0, s)$ and $\tilde{J}(s) = \min\{\hat{J}(s), \tilde{J}(s^+)\}$.

If $\tilde{J}(x^*) < \bar{c}$, then we extend the definition of $\tilde{J}$ to $[x^*, 1]$ by letting $\tilde{J} = h$. Function h is strictly increasing, and satisfies $h(x^*) = \tilde{J}((x^*)^-)$ and $h(1) = \bar{c}$. Moreover, $\int_{x^*}^1 h(x)dx > \bar{c}(1 - x^*) - \epsilon/2$. Notice that $\tilde{J}$ is strictly increasing and continuous on $(s, 1]$ (Property 1). In addition, $\tilde{J}(x) \leq \hat{J}(x)$ for $x \in [0, 1]$ (Property 2). Finally, $\int_0^1 \tilde{J}(x)dx > \int_0^{x^*} \hat{J}(x)dx + \int_{x^*}^1 \bar{c}dx - \epsilon > 1$ (Property 3). The right inequality in Property 3 follows from the choice of ϵ. Panel (b) in Figure 7, qualitatively illustrates the function $\tilde{J}(x)$ on the interval $(s, 1]$.

To complete the proof, we construct data D that is not justifiable. Consider the following data

$$D = \begin{cases} (s, a_c(d, s)) & c \in [0, \tilde{J}(s^+)] \\ \left(\tilde{J}^{-1}(c), a_c(d, \tilde{J}^{-1}(c))\right) & c \in R = (\tilde{J}(s^+), \tilde{J}(1)] \end{cases}$$

Property 1 implies that $\tilde{J}(x)$ is strictly increasing and continuous on $(s, 1]$, and therefore the inverse $\tilde{J}^{-1}(c)$ is well defined for all $c \in R$.

To see why data D is not justifiable, fix a client $c \in R = (\tilde{J}(s^+), \tilde{J}(1)]$. For the client c, the signal realization x_c is such that $c = \tilde{J}(x_c) \leq \hat{J}(x_c)$. The last inequality follows from Property 2. Therefore, the action associated with client c in data D is $a_c(d, x_c) = H$. In other words, any feasible density function J that justifies D must be weakly larger than $\tilde{J}(x)$ on $[0, 1]$. Thus,

$$\int_0^1 J(x)dx \geq \int_0^1 \tilde{J}(x)dx > \int_0^{x^*} \hat{J}(x)dx + \int_{x^*}^1 \bar{c}dx - \epsilon > 1$$

The last inequality follows from the Property 3. This inequality contradicts that the justification J is feasible and that the decision rule d is justifiable. □

Lemma 4. *Let $0 < \epsilon < \int_0^{x^*} \hat{J}(x)dx + \int_{x^*}^1 \bar{c}dx - 1$. Then there exists a strictly increasing function $\tilde{J}$ on $x \in \overline{V_+} = (s, x^*]$ such that*

1. $\tilde{J}(x)$ is strictly increasing and continuous on $(s, x^*]$ with $\tilde{J}(x^*) = \tilde{J}((x^*)^-)$,
2. $\tilde{J}(x) \leq \hat{J}(x)$ for all $x \in (s, x^*]$,
3. $\int_s^{x^*} \tilde{J}(x)dx > \int_s^{x^*} \hat{J}(x)dx - \epsilon/2$.

Proof. By Lemma 1, the function $\hat{J}$ is non-decreasing on $(s, x^*]$. Thus, it has at-most countable number of discontinuities on $(s, x^*]$. Let $\{I_n = (a_n, b_n)\}_{n \in \mathcal{N} \subset \mathbb{N}}$ be the countable collection of intervals such that $\hat{J}$ is continuous on (a_n, b_n), and is discontinuous at both the end-points.

Function $\tilde{J}$ is constructed on $(s, x^*]$ using the following procedure. For each $n \in \mathcal{N}$, consider the interval $I_n = [a_n, b_n]$. Define $\tilde{J}(a_n) = \hat{J}(a_n^-)$ if $a_n > s$. And, $\tilde{J}(b_n) = \hat{J}(b_n^-)$ if $b_n < x^*$. Moreover, let $\tilde{J}(x^*) = \tilde{J}((x^*)^-)$. For all x in the interior, choose $\tilde{J}$ to be any strictly increasing, continuous function such that $\tilde{J}(x) \leq \hat{J}(x)$ and $\int_{I_n} \tilde{J}(x)dx > \int_{I_n} \hat{J}(x)dx - \frac{\epsilon}{2^{n+1}}$.

Construction of $\tilde{J}$ implies that it satisfies Properties 1 and 2. To complete the proof, it suffices to show that Property 3 holds.

$$\int_s^{x^*} \tilde{J}(x)dx = \sum_{n \in \mathcal{N}} \int_{I_n} \tilde{J}(x)dx > \sum_{n \in \mathcal{N}} \int_{I_n} \hat{J}(x)dx - \sum_{n \in \mathbb{N}} \frac{\epsilon}{2^{n+1}} \qquad = \int_s^{x^*} \hat{J}(x)dx - \epsilon/2.$$

□

Lemma 5. *If the strategic justification $J^*(x) \leq \bar{c}$ for all $x \in [0, 1)$, then the expert's strategic justification is given by $J^*(x) = min\{f(x) + b + \lambda, \bar{c}\}$, where $-b \leq \lambda \leq 0$ and $\int_0^1 J^*(x)dx = 1$.*

Proof. The expert's problem is given by

$$\max_J \; -\int_0^1 \left[J_{\bar{c}}(x) \left(\frac{J_{\bar{c}}(x)}{2} - b \right) + (\bar{c} - J_{\bar{c}}(x)) f(x) \right] dx \tag{1}$$

$$\text{subject to } \; J \; \text{ is a non-decreasing, bounded, and atomless function,} \qquad \text{(monotonicity)}$$

$$\int_0^1 J(x)dx = 1, \qquad \text{(aggregate consistency)}$$

where $J_{\bar{c}}(x) = min\{J(x), \bar{c}\}$. If the strategic justification $J^*(x) \leq \bar{c}$ for all $x \in [0, 1)$, then adding the constraint $J(x) \leq \bar{c}$ to the expert's problem does not change the optimal solution. Moreover, it allows us to write $J_{\bar{c}}(x) = min\{J(x), \bar{c}\} = J(x)$ in the objective function. These modifications to the expert's problem reduce it to the problem (MP) with $\underline{x} = 0$, $\overline{x} = 1$ and $m = 1$. Thus, using the proof of Theorem 2, we obtain J^* proposed in the statement of the lemma. □

Lemma 6. *The objective function $H(J)$ is strictly convex in J.*

Proof. Fix a feasible density J. Observe that $H(J)$ is a linear function of $h(J(x))$. Therefore, it suffices to show that the $h(J(x))$ is strictly convex in $z = J(x)$. The proof is completed by observing that for any given x, the function $h(z) = z(\frac{z}{2} - b) + (\bar{c} - z)f(x)$ is strictly convex in z. □

Lemma 7. *Functionals H, e and $U = \int_{\underline{x}}^{\overline{x}} \eta_u(x)u(J)(x)dx$ are continuously Fréchet differentiable on $L^1 \cap B$.*

Proof. Recall Result 1 in Appendix B: ϕ is continuously Fréchet differentiable on $L^1 \cap B$ if ϕ is Fréchet differentiable at every J. Thus, it suffices to show that functionals mentioned in the statement of the lemma are Fréchet differentiable at every $J \in L^1 \cap B$.

Consider the functional $H : L^1 \cap B \to \mathbb{R}$ and fix $J \in L^1 \cap B$. Let $g(x) = J(x) - f(x) - b$. Since $J(x), f(x) \leq M$ for some $M > 0$, function $g(x) \in L^\infty$. By Result 2 in the mathematical preliminaries, function g defines a bounded linear functional $H'_J(\phi) = \int_{\underline{x}}^{\overline{x}} g(x)\phi(x)dx$. Below, we establish that H'_J is the Fréchet derivative of H at J.

$$\begin{aligned}\lim_{t\to 0} \frac{|H(J+t) - H(J) - H'_J(t)|}{\|t\|_1} &= \lim_{t\to 0} \frac{1}{\|t\|_1} \left| \int_{\underline{x}}^{\overline{x}} [h(J+t) - h(J) + (J(x) - f(x) - b)]\, t(x)dx \right| \\ &= \lim_{t\to 0} \frac{1}{\|t\|_1} \left| \int_{\underline{x}}^{\overline{x}} [t(x)/2]\, t(x)dx \right|. \\ &\leq_{(a)} \lim_{t\to 0} \frac{\|t\|_1 \|t\|_\infty}{2\|t\|_1} = \lim_{t\to 0} \frac{\|t\|_\infty}{2} = 0.\end{aligned}$$

Where inequality (a) is obtained by Hölder's inequality. As t approaches 0 in L^1, the Lebesgue measure of points where $t(x)$ takes non-zero values also approaches to 0. In other words, $\|t\|_\infty = \inf\{M : \mu\{x : t(x) > M\} = 0\}$ approaches 0. Thus, H is Fréchet differentiable on $L^1 \cap B$.

In a similar manner, functionals e and U are continuously Fréchet differentiable on $L^1 \cap B$. In particular, $e'_J(\phi) = \int_{\underline{x}}^{\overline{x}} \phi(x)dx$ and $U'{}_J(\phi) = \int_{\underline{x}}^{\overline{x}} \eta_u(x)\phi(x)dx$. □

Lemma 8. *Function J^* proposed in the statement of Theorem 2 satisfies regularity conditions.*

Proof. The following result (Lemma 1.14 on Page 85 of Hinze et al. (2008)) states the sufficient conditions for J^* to satisfy regularity conditions.

Lemma. *J^* is a regular point if $e'_{J^*} : L^1 \to \mathbb{R}$ is surjective and if there exists $\tilde{J} \in L^1 \cap B$ such that*

$$e'_{J^*}(\tilde{J} - J^*) =_{(a)} 0,$$
$$u(J^*) + u'{}_{J^*}(\tilde{J} - J^*) \in_{(b)} int(K),$$

where $K = \{v \in L^2 : v(x) \leq 0\}$.

By the last line in the proof of Lemma 7, the Fréchet derivative $e'_{J^*}(\phi) = -\int_{\underline{x}}^{\overline{x}} \phi(x)dx$. Fix $r \in \mathbb{R}$ and let $\phi = -(\overline{x} - \underline{x})r$. Since $e'_{J^*}(\phi) = r$ the requirement of surjectivity is satisfied.

Let $\tilde{J} \in L^1 \cap B$ be a function such that $\int_{\underline{x}}^{\overline{x}} \tilde{J}(x)dx = m$ and $\tilde{J}(x) < \overline{c}$ for all $x \in [\underline{x}, \overline{x}]$. This function satisfies condition (a). Since $u'{}_{J^*}(\phi) = \phi$, the left hand side of equation (b) becomes $J^*(x) - \overline{c} + \tilde{J}(x) - J^*(x) < 0$. Therefore, equation (b) is also satisfied.

To complete the proof, it suffices to show that the function $\tilde{J}$ with the proposed properties exists. Since $m < \int_{\underline{x}}^{\overline{x}} \min\{f(x)+b, \overline{c}\}dx$, we have $m < \overline{c}(\overline{x} - \underline{x})$. Thus, the constant function $\tilde{J} = \frac{m}{\overline{x}-\underline{x}}$ satisfies the proposed properties. □

Lemma 9. *J^* is a unique solution to Problem (MP).*

Proof. Since J^* is a solution to the Problem (MP), there exists a $\epsilon > 0$ such that the following statement holds: For any $\tilde{J} \in \mathcal{Y}$ that satisfies $\|J - \tilde{J}\|_1 < \epsilon$, we have $H(J^*) \leq H(\tilde{J})$.

Suppose $\hat{J}$ that differs from J^* on $[\underline{x}, \overline{x}]$ is also a solution to the Problem (MP). Thus $H(J^*) = H(\hat{J})$. Consider the function $\tilde{J} = (1-\mu)J^* + \mu\hat{J}$, where $\mu > 0$ is such that $\mu\|J^* - \hat{J}\|_1 < \epsilon$. Since the constraint set $\mathcal{Y}$ is convex, function $\tilde{J} \in \mathcal{Y}$. In addition, $\tilde{J}$ satisfies $\|J^* - \tilde{J}\|_1 < \epsilon$. By the strict convexity of H in J (Lemma 6), we obtain the inequality $H(\tilde{J}) < (1-\mu)H(J^*) + \mu H(\hat{J}) = H(J^*)$. This latter inequality contradicts to the fact that J^* is a solution. Thus, there does not exists $\hat{J} \neq J^*$ that is also a solution to the Problem (MP). □

Lemma 10. *Let $m = 1 - \int_{x_j}^{1} \underline{f}_{x^*}(x)dx$. If the strategic justification $J^*(x) \leq \overline{c}$ for all $x \in [0, x_j)$, then the expert's strategic justification is given by $J^*(x) = min\{f(x)+b+\lambda, \overline{f}(x), \overline{c}\}$, where $-b \leq \lambda \leq 0$ and $\int_0^{x_j} J^*(x)dx = m$.*

Proof. Consider the expert's problem

$$\max_{J} \; -\int_0^{x_j} \left[J_{\overline{c}}(x) \left(\frac{J_{\overline{c}}(x)}{2} - b \right) + (\overline{c} - J_{\overline{c}}(x)) f(x) \right] dx \tag{P-A}$$

$$\begin{aligned}
\text{subject to} \quad & J \text{ is a non-decreasing, bounded \& atomless function,} && \text{(monotonicity)} \\
& \underline{f}(x) \leq J(x) \leq \overline{f}(x) \quad \text{for } x \in [0, x_j), && \text{(local consistency)} \\
& \int_0^{x_j} J(x)dx = m, && \text{(aggregate consistency)}
\end{aligned}$$

where $J_{\overline{c}}(x) = min\{J(x), \overline{c}\}$.

Since the optimal solution $J^*(x) \leq \overline{c}$ for all $x \in [0, x_j)$, replacing the affine constraint $J(x) \leq \overline{f}(x)$ with the affine constraint $J(x) \leq \min\{\overline{f}(x), \overline{c}\}$ does not change the optimal solution. Moreover, it allows us to write $J_{\overline{c}}(x) = min\{J(x), \overline{c}\} = J(x)$ in the objective function. Before solving this modified problem, we relax the constraint $J(x) \geq \underline{f}(x)$. The resulting problem is analogous to the problem (MP) with the difference that the affine constraint is now changed to $J(x) \leq \min\{\overline{f}(x), \overline{c}\}$. Moreover, $\underline{x} = 0$, $\overline{x} = x_j$, and $m = 1 - \int_{x_j}^{1} \underline{f}_{x^*}(x)dx$. Thus, using the proof of Theorem 2, we obtain J^* on the interval $[0, x_j)$. To complete the proof of Lemma 10, observe that the proposed J^* satisfies $J^*(x) \geq f(x) \geq \underline{f}(x)$ for $x \in [0, x_j)$. □

References

Alonso, Ricardo, and Niko Matouschek. 2008. "Optimal Delegation." *The Review of Economic Studies*, 75(1): 259–293.

Bhattacharya, Vivek, Gaston Illanes, and Manisha Padi. 2020. "Fiduciary Duty and the Market for Financial Advice." *Working Paper.*

Cherepanov, Vadim, Timothy Feddersen, and Alvaro Sandroni. 2013. "Rationalization." *Theoretical Economics*, 8(3): 775–800.

Crawford, Vincent P, and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica: Journal of the Econometric Society*, 1431–1451.

Eliaz, Kfir, Ran Spiegler, and Heidi C. Thysen. 2021. "Strategic Interpretations." *Journal of Economic Theory*, 192: 105192.

Frankel, Alex. 2016. "Delegating Multiple Decisions." *American economic journal: Microeconomics*, 8(4): 16–53.

Frankel, Alexander. 2014. "Aligned Delegation." *American Economic Review*, 104(1): 66–83.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *The Journal of Law and Economics*, 24(3): 461–483.

Hinze, Michael, René Pinnau, Michael Ulbrich, and Stefan Ulbrich. 2008. *Optimization with PDE constraints.* Vol. 23, Springer Science & Business Media.

Holmstrom, Bengt Robert. 1978. *On Incentives and Control in Organizations.* Stanford University.

Jackson, Matthew O, and Hugo F Sonnenschein. 2007. "Overcoming Incentive Constraints by Linking Decisions." *Econometrica*, 75(1): 241–257.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Lehrer, Ehud, and Roee Teper. 2011. "Justifiable Preferences." *Journal of Economic Theory*, 146(2): 762–774.

Melumad, Nahum D, and Toshiyuki Shibano. 1991. "Communication in Settings with No Transfers." *The RAND Journal of Economics*, 173–198.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics*, 12(2): 380–391.

Ok, Efe A. 2011. *Real Analysis with Economic Applications.* Princeton University Press.

Rayo, Luis, and Ilya Segal. 2010. "Optimal Information Disclosure." *Journal of political Economy*, 118(5): 949–987.

Royden, Halsey Lawrence, and Patrick Fitzpatrick. 1988. *Real Analysis.* Vol. 32, Macmillan New York.

Schwartzstein, Joshua, and Adi Sunderam. 2021. "Using Models to Persuade." *American Economic Review*, 111(1): 276–323.

Spiegler, Ran. 2002. "Equilibrium in Justifiable Strategies: A model of Reason-based Choice in Extensive-form Games." *The Review of Economic Studies*, 69(3): 691–706.